


How do different types of carbon labels influence consumer preferences?

The effect of brand level information versus product level information.

<i>Student:</i>	Remco Wezepoel
<i>Registration number:</i>	910110948110
<i>Chair group:</i>	Economics of Consumers and Households
<i>First supervisor:</i>	dr. M.J.J. Handgraaf
<i>Second supervisor:</i>	dr. J.A.C. van Ophem
<i>Course code:</i>	ECH-80433
<i>Date:</i>	9-5-2016
	This label indicates whether or not this thesis was written with special attention regarding to climate change

Preface and acknowledgements

This MSc thesis has been written as a part of my Master's degree "Management, Economics and Consumer Studies" at the Wageningen University. During my studies, I developed an interest for labelling on products. However, our study almost always dealt with labels on food products. I think it is fascinating to see how labels can influence consumer's thoughts, preferences and behaviour, but I did not want to reinvent the wheel. Searching for an original perspective regarding product labels, I found an interesting article about the effect of brand names, which was really attracting me. For current research, both topics are combined, in order to compare the effect on consumer preferences regarding four carbon labels containing information on a brand level or on a product level.

I would like to express my gratefulness to dr. Michel Handgraaf, for being understanding towards my situation and for not demanding to hand in incomplete work. But moreover, for his guidance, patience, support, and positive feedback throughout this research. His in-depth knowledge on product labelling, consumer behaviour and statistical analysis, together with his clear and practical feedback, have been of great value during this research. The process of writing this thesis and performing extensive research regarding consumer behaviour gave me the opportunity to discover new strong points and weaknesses. Furthermore, it gave me a better insight in what I like and do not like to do, which will help me in my quest to find a job after graduating. Finally, I would like to thank dr. Johan van Ophem for being second supervisor of this report. Also the ECH chair group, my fellow students, friends and family have each supported me a lot in their own way.

By handing in this report, my time as a student comes to an end. I am happy to announce that I was offered a job at L'Oréal, which I have accepted. The last six years of studying have paid off and I am ready to start my first real fulltime job in May 2016. I am looking forward to start this new adventure!

Remco Wezepoel

Wageningen / The Hague, May, 2016

Table of contents

Preface and acknowledgements	i
Table of contents	ii
Executive summary	iv
Introduction	1
Labelling	4
Product versus brand level.....	9
Moderating variables	12
New Environmental Paradigm	12
Social Norms	14
Perceptions of influence	16
Method.....	18
Results.....	24
WTP analysis	28
WTP Brand versus Product level, positive labels.....	28
WTP Brand versus Product level, negative label	30
WTP Brand level, positive versus negative labels.....	32
WTP Product level, positive versus negative labels.....	34
Conclusion WTP	37
Liking analysis.....	39
Liking Brand versus Product level, positive label	39
Liking Brand versus Product level, negative label.....	42
Liking Brand level, positive versus negative labels	44
Liking Product level, positive versus negative labels	47
Conclusion liking.....	50
Attractiveness analysis.....	52
Attractiveness Brand versus Product level, positive label	52
Attractiveness Brand versus Product level, negative label.....	54
Attractiveness Brand level, positive versus negative labels	57
Attractiveness product level, positive versus negative label.....	59
Conclusion attractiveness.....	63
Recall analysis.....	65
Recall exact brand versus product level, positive label	65

Recall approximately, brand versus product level, positive label	68
Recall exact brand versus product level, negative label	70
Recall approximately brand versus product level, negative label.....	71
Recall exact positive versus negative labels, Brand level	73
Recall approximately positive versus negative label, Brand level	75
Recall exact positive versus negative label, Product level	77
Recall approximately positive versus negative label, Product level	80
Conclusion recall.....	82
Assessment label test	83
Discussion.....	84
Limitations.....	92
Conclusion	94
Recommendations for future research	95
Literature.....	97
Appendices	106

Executive summary

Purpose - Carbon dioxide is a very important polluting greenhouse gas, arising mainly from human behaviour. Reducing carbon emission is attempted to reach by means of carbon emission labels on products. Current research was conducted in order to determine the effect of brand level information versus product level information on consumer preferences regarding 4 different types of carbon emission labels.

Design/methodology - Current study used a 4 (type of label: checkmark, traffic light, black score, colored score) x 2 (information level: brand versus product) mixed model design. A total of 131 respondents between 18 and 37 years participated in an online survey that began with questions regarding their New Ecological Paradigm score (NEP), perceptions of influence, and belief in social norms. Hereafter, the respondents saw one out of the four carbon labels (on both a brand level and on a product level), where after questions regarding their willingness to pay (WTP), ratings of liking and rating of attractiveness were asked. The survey ended with a recall check regarding the displayed label and with some questions regarding assessment of the labels.

Findings – Results of current study show that respondents were willing to pay more for a fridge with a carbon label on a product level, compared to a brand level. Also, positive carbon labels on a product level were preferred over positive carbon labels on a brand level. NEP appeared to have a positive effect on WTP for positive carbon labels. Type of label (checkmark, traffic light, black score, and colored score) does not influence participants' WTP and rates of attractiveness. However, rates of liking are influenced by type of label, in a way that a positive checkmark and traffic light are preferred over a positive black score and a positive colored score. If a label was negative, the black score was preferred over the traffic light and the colored score.

Positive carbon labels on a brand level are significantly better remembered compared to positive carbon labels on a product level. Positive carbon labels on a brand level are better remembered compared to negative carbon labels on a brand level. If a carbon label is positive, the checkmark and the traffic light label are both significantly better remembered compared to the black score labels and the colored score labels.

Finally, the traffic light label was indicated as best interpretable.

Research limitations – The duration of the used survey was a limitation which has led to many incomplete surveys and maybe to not-seriously completing the survey. Furthermore, the main limitation of this study is that the survey was not fully built based on the hypotheses. As a result, not all hypotheses were able to be measured.

Research implications – This research ends with recommendations for future research, in which potential implications regarding E-mail marketing and framing product information for environmental concerned companies/brands are discussed.

Introduction

Carbon dioxide is the most important polluting greenhouse gas (GHG) arising mainly from human behavior (Kellogg and Schware, 1981). It is emitted through activities like agricultural-, industrial- and transportation operations (Kellogg, 1979). Increased emission of carbon dioxide influences the climate by absorbing parts of the infrared radiation which is emitted by the surface of the Earth. As a result, global temperatures have continued to rise and fall during the past decades (Crowley, 2000).

Increased concern about climate change has stimulated interest regarding greenhouse gas emissions since the late 1980s (Kolk and Pinkse, 2004). Consumers as well as organizations are more and more interested in the emission of greenhouse gases during production, processing, retailing and use of consumer goods (Edwards-Jones et al., 2009; Kolk, Levy and Pinkse, 2008). However, motives differ. Whereas organizations aim to use a sustainable image in order to show responsibility and leadership in the climate change issue (Carbon Trust, 2008), consumers indicate that they have personal health reasons and/or concern for the environment (Röös and Tjärnemo, 2011).

In order to stimulate consumers in making informed choices, many marketing communication options exist. Examples are websites, providing information at the point of purchase, or, an option which is often being discussed lately, by using a label on the product packaging (Olofdotter and Juul, 2008). With regard to GHG emissions, a so called "carbon label" could be an example of latter option (Röös and Tjärnemo, 2011).

Many variations among these labels exist. There are traffic light colours (Sacks, Rayner and Swinburn, 2009) and pictures of a footprint with messages like "Lower/medium/higher CO₂" (Vanclay et al., 2011). Furthermore, the labels could consist of a logo which indicates that a product is produced with special consideration taken to reduce emissions (Röös and Tjärnemo, 2011), or it can contain calculations of product carbon footprints, which can indicate the amount of GHG produced during different stages of the life cycle of products and services like production, processing, transportation, use, and disposal (Bowling and Gibbon, 2009).

Information regarding carbon dioxide on labels is expected to help consumers by increasing the salience of environmental sustainability (Upham and Bleda, 2009). It also aims to make consumers aware of how their product- or used service choice affects the emissions of GHGs (Kimura et al., 2010). In the end, this should stimulate consumers to contribute to the fight against climate change by consciously selecting their consumer goods (Kimura et al., 2010). However, the effect of (carbon) product labels remains

unclear. Upham and Bleda (2009) state that labels may confuse consumers or might be ignored, whereas Kimura and colleagues (2010) found that carbon footprint labels contribute to creating value for consumers with regard to food products. However, mere exposure to the footprint information seemed not to be sufficient to achieve understanding of the information among consumers (Kimura et al., 2010).

There seem to be many studies reporting that consumers are rarely affected by claims of sustainability (Valor, 2008; Weightman and McDonagh, 2004), whereas more recent studies show that consumers are willing to pay more for products labelled as “green products” (Young et al., 2010). Research also shows that consumers are concerned about climate change and interested in carbon labelling (Vanclay et al., 2011).

Even though there is evidence showing that consumer interest in “green products” has increased (Cohen and Vandenberg, 2012), opposing results remain with regard to the actual purchase behavior of these products. This raises questions regarding the effectiveness of (carbon) product labels. Consumers tend to claim a large interest in carbon labelled food (Röös and Tjärnemo, 2011), but actual consumer behavior often lacks performance of the ethical intended behavior (Carrington, Neville and Whitwell, 2010). This gap between intentions and behavior can have several reasons. Examples can be label confusion and perceived social control (Sheeran, 2002), which implies that the label gives too much, too detailed or too vague information, or that consumer behavior depends on other people’s approval or disapproval. Furthermore, low perceived quality of a “green” product and the price of this product are mentioned as variables which influence the gap between intentions and behavior (Griskevicius, Tybur, and van den Bergh, 2010).

Consumer often rely on brands. Hoyer and Brown (1990) showed that brand awareness leads to a convenient automatic choice towards products, even when the quality of the product with the “known” brand is lower as compared to other brands in the choice set. Furthermore, consumers tend to evaluate fewer alternatives when there is at least one brand in the choice set which they are familiar with (Hoyer and Brown, 1990). Hoyer (1984) found that consumers tend to spend minimal time and cognitive effort in low costs purchase situations. In these situations, heuristics are frequently used. This implies that consumers may choose a product based on simple shortcuts like prices, packaging, or brand (MacKenzie and Spreng, 1992). However, situations which involve more costs in terms of time and money, so called “high-cost situations” (Diekmann and Preisendörfer, 2003), require more rational behavior, which leads to elaborated decisions (Zintl, 1989).

Examples of high-cost situations are changing transportation behavior from travelling by car to taking public transportation, or purchasing a new TV. In high-cost situations, consumers extensively compare and evaluate different brands (Kressmann et al., 2006). This shows the importance of brands in both high- and low-cost situations.

Product packages contain a lot of information. Further increasing the amount of information on packages may pose a problem, since this might increase the possibility of information overload (Kimura et al., 2010). Decision quality is negatively influenced by information overload and the overload of information might even lead to consumers totally ignoring the information (Eppler and Mengis, 2004; Hwang and Lin, 1999). However, brand information fits the heuristics which consumers tend to use in purchase decisions. As a result, we will investigate how consumers respond to carbon labels with information on a brand level versus on a product level.

This all points out that deeper insight is needed with regard to consumer evaluation of different types of carbon labels combined with how these labels influence their decision making. Currently, **Climate-KIC** is developing a new carbon label. This will be a single score between zero and 100, indicating the carbon friendliness of a product. In order to provide this public-private innovation partnership recommendations regarding this carbon label, this study will investigate how four different types of carbon labels influence consumer willingness to pay, satisfaction, and product preference. A distinction will be made between the four types of labels providing information on a brand level versus on a product level. Finally, recall will be measured in order to see whether indicated preferred labels are also the best remembered labels. This leads to the following research questions:

- 1. What is the effect of different types of carbon labels on consumers' preferences regarding willingness to pay, attractiveness and liking?*
- 2. How does information on a carbon label on a brand level versus on a product level moderate the effect of different types of labels on consumer's preferences?*
- 3. Does level of information (brand versus product) have an effect on recall of the label?*

Labelling

Emerging evidence shows a remarkable paradox; there is proof that society is increasingly concerned about the environment, while market success of environmental friendly products stays behind (Kalafatis, Pollard, East and Tsogas, 1999). Even though this claim might be a bit outdated, market success of environmental friendly products is still small. Possible explanations can be found in the literature. Pickett-Baker and Ozaki (2008) found that consumers felt good about buying environmental friendly products, but these consumers indicated that they had difficulties with identifying these products. In line with these findings, Rand Corporation (2004) showed that manufacturers of environmentally-improved products make little or no green claims in their advertising.

Labelling could be a strategy for manufacturers to communicate this information about environmental friendliness or other product features such as healthiness, energy usage, ozone-friendliness and used material of the packaging to consumers (D'Souza, Taghian and Lamb, 2006; Wogalter and Young, 1994). Eco labels like carbon labels are based on persuasive communication. This implies that the information aims to change consumer behavior (Gössling and Buckley, 2014). Even though a large amount of research has been conducted regarding product labels and consumer behavior, no general consensus has been found yet. Chase and Smith (1992) found that a large majority of consumers' purchase decisions were influenced by messages in advertising and product labelling. More recent research has shown that labels may confuse consumers or may even be ignored (Upham and Bleda, 2009), whereas Young and colleagues (2010) found that consumers are willing to pay more for products labelled as "green products".

Many more contradicting results regarding carbon labelling are found in the literature. Hartikainen and colleagues (2014) showed that 909 out of 1010 respondents stated that carbon labels would have at least a little impact on their buying behavior with regard to food products. However, this was only the case after other purchasing criteria such as price and taste were met. Another example supporting the role of carbon labels is described in the paper of Cohen and Vandenberg (2012) about the potential of carbon labelling, which concludes that there is sufficient evidence that purchase decisions are directly or indirectly influenced by claims about environmental benefits. On the contrary, Upham, Dendler and Bleda (2011) found that consumers encountered difficulties in making sense of carbon labelled grocery products. Moreover, they found that there was little evidence for carbon labelling mediating in product selection.

The literature provides different reasons explaining why product labels might not be effective. Misunderstanding of the meaning of information on the label is a substantial reason (D'Souza, Taghian and Lamb, 2006; Hartikainen et al., 2014). Moreover, perceived high product price, accustomed purchase habits, lack of marketing and information, lack of trust in the labelling system, and low perceived consumer effectiveness are mentioned as causes of label ineffectiveness (Röös and Tjärnemo, 2011). It has also been mentioned that product labels contain too much information, and that too many labels exist, which confuses consumers or make them even ignore the labels completely (Upham and Bleda, 2009). Besides that, too many labels exist, which makes it nearly impossible for consumers to understand all the different meanings (Hawthorne, Moreland, Griffin, and Abrams, 2006).

Label design

A recent paper by Cohen and Viscusi (2012) provides recommendations related to carbon footprint labels. In their paper, Cohen and Viscusi state that carbon labels focus more on externalities which will probably not be apparent until many decades after the purchase, rather than the direct costs and benefits for individuals which are generally emphasized by product labels. In order to be effective, Cohen and Viscusi (2012) state that consumers must “receive the information, process the information, believe the information and use it to update beliefs that potentially can influence decisions regarding the product”. Labels which are placed on products make sure that consumers receive the information, as long as the label does not contain more than four or five informational parts, since consumers generally cannot process higher amounts of information (Cohen and Viscusi, 2012). Furthermore, Cohen and Viscusi state that content, format and usage of the carbon label matters. A broad audience has to be able to understand the information on the label, and the label should be sufficiently large to be noticed and readable. Uniformity in the index, which is used to show the carbon footprint of a product is desirable, and the use of labelling should be selective, in order to avoid information overload. If every product category contains a carbon label, it will be hard for consumers to make judgements about when the label is important and when it is not. Likewise, the carbon labels should convey a consistent mechanism for communicating the environmental effects, in order to convey a meaningful message (Cohen and Viscusi, 2012).

Finally, targeting of specific product categories appeared to be important in order to create a successful label. This implies that product categories should wear the same label, in order to avoid confusion and distrust. It is also important that a label displays clarity, which results in better understanding and increased willingness to cope with the information on the label (Truffer et al., 2001). Finally, Cohen and Vandenberg (2012) emphasize the importance of additional information next to the statistical information on a label. The purpose of this additional information is to communicate towards consumers how their decisions have an impact on carbon emission.

Labelling in the literature

Carbon labelling is a relatively new subject in the literature. However, in terms of design, there are many similarities with food packaging labels, which has been investigated more intensively. This makes literature regarding labels on food packages relevant as a benchmark. Three different types of labels lead the debate on labels on product packages (van Herpen, Seiss and van Trijp, 2012). Even though their reasoning applies to the area of nutrition labelling, the labels fit the description given by Cohen and Viscusi (2012). The three mentioned basic types of labels are the so called signpost logos (e.g., health tick, conscious choices logo), multiple traffic light (MTL) labels, and labels communicating guidelines daily amounts (GDA) in terms of numbers. Where the signpost logos exist of an overall “approval mark” like a single checkmark or a thumb up, MTL labels use colors in order communicate a message with regard to low, medium, or high levels e.g. healthiness. GDA labels provide numerical details (van Herpen et al., 2012).

For experiments with respect to carbon labelling and consumer perceptions, many types of labels are used. A frequently used label within this field is the multiple traffic light label, as well as a single number which shows information about emissions in grams, a carbon footprint picture which displays an amount of CO₂, a single check mark, and a label with a scale indicating environmental impacts of the product are commonly used (for examples, see: Banerjee and Solomon, 2003; Kimura et al., 2010; Sacks, Rayner and Swinburn 2009; Vanclay et al., 2011). Banerjee and Solomon (2003) evaluated consumer and marketer responses towards five US eco-labels and concluded that a single “Energy Star” label was most effective, together with other simple logos (like a checkmark and a scale) as compared to more complex information-disclosure labels, since the standardized labels appeared to be easy to process by consumers. Based

hereon, we assume that simple, easy to understand labels are more effective than more complex labels which contain a lot of information. As a result, the current study will focus on these so called “simple labels”. Frequently used labels and labels which consumers are familiar with are the check mark label and the multiple traffic light label. Furthermore, as mentioned before, a new carbon label is being developed by Climate-KIC . This label will be a single score between zero and 100, indicating the carbon friendliness of a product. However, the authors of current study are not convinced that a single score will convince consumers of its meaning, a fourth label will be added. Since we assume that color also influences decisions (as proven with the multiple traffic light label), the fourth label will be a score between zero and 100 combined with three different colors. The scores from zero to 33 will be red, scores from 34 to 66 will be orange, and scores from 67 to 100 will be green. This addition is based on recommendations of a focus group in a study performed by Upham, Dendler, and Bleda (2011). In this study, participants of the focus group stated clearly that the use of a simple label with a “color-coded traffic light style system” would make the information more meaningful to them.

Attractiveness of the label vs. effectiveness of the label

With regard to “simple labels”, many research has been conducted in the field of nutritional products. Research of van Herpen and van Trijp (2011) has shown that traffic light labels and single logo labels (a green tick) enhance healthy product choices. Even while consumers rated a nutrition value table (numerical label) more positive as compared to a traffic light label and a green tick logo, the numerical label appeared to obtain little attention and did not stimulate healthy choices. As a result, we expect that consumers will prefer the single logo label (checkmark and traffic light) over a score label (black or “traffic light” colored).

More evidence supporting the preference of the traffic light and single logo labels over a score label can be found in another study from van Herpen, Hieke, and van Trijp (2014). This study shows that labels with a reference point such as a traffic light and a check mark label are more effective in consumers’ product evaluation compared to labels lacking reference point information. A numerical score is an example of a label lacking a reference point and hence we assume that the traffic light label and the single logo label (checkmark) will have a higher influence on satisfaction as compared to the score labels. Finally, with regard to labelling, research of Strasser and colleagues (2012) showed a difference in recall of graphic labels and text-only labels. Results showed that graphic

labels were significantly better recalled compared to text-only labels. Even though these findings concern warning labels and texts against smoking, it is plausible to generalize these findings towards carbon emission, since these labels are also meant as a “warning” about pollution. Based hereon, we assume that a graphic label (checkmark or traffic light) will be better recalled than a score label (black or colored).

Drawing on this line of results in the field of labels regarding nutrition and smoking, we assume the following hypotheses:

H1: Carbon labels will be better understood when the label is a traffic light label or a checkmark label as compared to a score label (black or colored).

H2: Positive labels (green checkmark, green traffic light, a black score above 66 and a green score above 66) will be preferred over negative labels (red checkmark, red traffic light, a black score under 33 and a red score under 33).

H3: Carbon emission labels containing a checkmark and a traffic light label will be preferred over a score label (black or colored).

H4: The checkmark and the traffic light labels are better recalled as compared to a black score and a colored score label.

H5: A colored score label will be recalled better as compared to a black score label.

Product versus brand level

When consumers make purchase decisions, different alternatives are evaluated. Research has shown that important evaluation criteria are price, brand and store name (Dodds, Monroe and Grewal, 1991). This implies that a brand (of the store as well as of the product) can affect consumer decision making. The literature is rich in studies regarding brands and how brands can have an influence on consumers. It is known that brand familiarity has an influence on information search. A good example is observed in a study of Biswas (1992), which showed that consumers who are shopping for a familiar brand intend to spend less time on the process as compared to consumers who are shopping for an unfamiliar brand. Brand names alter consumers' judgements about products and their quality. Research of Dodds and Monroe (1985) showed that adding a brand name significantly increased participants' perceived quality and willingness to buy regarding household products. Furthermore, brand names can help consumers in making a final purchase decision while comparing several products (Hillenbrand, Alcauter, Cervantes and Barrios, 2013) and enhances consumer purchase intentions among certain product categories (Arora and Stoner, 1996). Furthermore, Persson (2010) states that consumers prefer buying "a brand" over buying "a plain product". As a result, we assume that consumers are willing to pay more for a "brand" than for a "plain product" when a label is positive. This implies that we assume that brand information has more impact than product level information. As a result, we also assume that a negative label on a brand level leads to lower willingness to pay compared to a negative label on a product level.

Keller (1993) shows that consumers will make associations with a brand because of the image they have of the brand. This will occur through the process of categorization.

A similar process occurs when evaluating a product and its attributes. Consumers can associate a brand or a product with a particular feature, usage situation or person (John, Loken, Kim, and Monga, 2006). The distinction between products and brands will most likely lead to different associations in the mind of consumers due to the process of categorization. Categorisation is an important term in the evaluation of products and brands. When a consumer perceives an object (product or brand in this case), the brain will look for similarities and differences in such a way that the perceived object can be placed in a category. Categorization processes are used to store and retrieve information in the memory and to make sense of events and objects (Schoormans and Robben,

1997). Through these processes, purchase decisions are guided by the associations that are evoked by the stimuli (Aaker, 1991). We assume that brands will evoke different associations and feelings compared to products. More precisely, Aaker (1996) states that the result of the associative network theory, which is an associative network, contributes to creating a value and an image about the product/brand to consumers. Brands will evoke stronger associations compared to products because products often have a limited amount of associations whereas brands are often more of an umbrella term for many different products and therefore associations. Continuing within this line of reasoning, it seems that it is often hard to separate the link between products and brands, but it is also logical that consumers will evaluate aspects of a product and of a brand differently. There might be a difference in how consumers evaluate e.g. a “green brand” versus a “green product” or a regular or non-green brand. This becomes more complex when people receive information on a product level and on a brand level, where it is possible that the information might seem contradictory. An example could be where a brand is green but the product is not or the other way around. It might be weird if a brand like Hummer would introduce a hybrid model, right?

Information of a product and its attributes is often provided to the consumer via a label. Research described in the currently available literature however, does not make any distinction between whether a label is on a brand level or on a product level. Looking at examples of the application of labels in the “real world”, labels are almost always on a product level. This might be seen as odd, since brands evoke associations of for example quality and environmental friendliness (Keller, 1993). However, products and brands are highly related. A brand can contain several products. If a product of brand X contains a carbon friendly label, this does not imply that other products of the same brand X also are environmentally friendly. Literature makes a distinction in how consumers view a product or a brand. For example, when consumers think of a brand name like Apple, this immediately raises all kinds of associations and a whole array of products, while thinking of an iPhone directly creates a link with for example Apple, expensiveness, or with mobile phones. It seems that consumers differentiate products from brands and evaluate them in different ways. However, little research in the field of how this process works has been conducted. As stated before, consumers are willing to pay more for a “brand” compared to a “plain product”, therefore we assume that consumers are willing to pay more for a product with a label on a brand level compared to a label with information on a product level, when the label is positive.

The usage of positive labels is often observed as a way to depict the levels of carbon emission. It could be interesting to make a distinction between information about carbon emission on a brand level versus on a product level. The fact that brands influence consumers in creating purchase intentions and making purchase decisions might imply that a carbon label on a brand level can help consumers in making evaluations and decisions more easily. Research from Pickett-Baker and Ozaki (2008) showed that making use of brands to promote products which are environmentally friendly could increase sales of these products. Respondents in their study stated that they were more likely to purchase certain brands if they knew that these companies manufactured products in an environmentally friendly way. As a result, we assume that carbon labels on a brand level are preferred over carbon labels with information on a product level.

Current research uses the distinction between products and brands, in order to find out whether a carbon information label has different effects on consumers' perceptions about certain products on a brand level or on a product level.

However, since brands can have a lot of influence (think about brand communities; Apple, Starbucks, Volkswagen), we assume that negative labels on a brand level are seen as more negatively compared to negative labels on a product level. As a result, we assume that consumers are willing to pay less for a product with a negative label on a brand level, compared to a negative label on a product level. This would imply that labels which provide carbon emission information on a brand level would have a strong influence on consumers. However, we use unknown brands which would evoke little associations. As a result, we assume that the label with information on a product level will have less influence on consumer ratings regarding that product as compared to a label on a brand level. Taking this together, it leads us to the following hypotheses:

H6: Respondents are willing to pay more for a positive label on a brand level compared to a positive label on a product level.

H7: Respondents are willing to pay more for a negative label on a product level compared to a negative label on a brand level.

H8: Respondents will give higher ratings of liking and attractiveness for labels on a brand level compared to labels on a product level.

Moderating variables

In previous chapter, main effects of type of label and brand level versus product level are discussed. It is likely that expected effects are stronger or weaker for certain consumers with specific individual characteristics. As a result, this chapter explains the constructs “New Environmental Paradigm”, social norms, and perceptions of influence. These constructs serve as independent moderator variables in this research. Assumptions will be drawn upon the literature and hypotheses will be stated, since it is expected that high or low scores on these constructs lead to different scores on WTP, liking, attractiveness, and recall.

New Environmental Paradigm

Since current research deals with carbon labels, it is important to check whether participants are high- or low involved with environmental concerns. Regarding measures of environmental attitudes, only three scales (the Ecology Scale, the Environmental Concern Scale, and the New Environmental Paradigm (NEP) Scale) have been widely used (Hawcroft and Milfont, 2010). Dunlap and Jones (2003) state that items of the Ecology Scale and the Environmental Concern Scale have become outdated as a result of new environmental issues which have been emerged since they measure specific topics. NEP Scale measures general beliefs about pro-environmental orientation, and thus avoids this issue of becoming outdated. This makes the NEP Scale suited for current research.

Re-examination of this NEP Scale (originally published in 1978 by Dunlap and van Liere) showed some points which were affecting the scale’s reliability. Mentioned shortcomings were a lack of internal consistency, use of “old fashioned” language and poor correlation between the scale and behavior (Geller and Lasley, 1985; Stern, Dietz, and Guagnano, 1995). Dunlop and colleagues (2000) responded to the criticism by revising the scale, and thus improving the original one which made the scale more reliable (Hawcroft and Milfont, 2010).

The NEP scale which was used in current research consisted of fifteen statements, measuring the environmental concern of participants. A high score on the scale implies high environmental concern, while a low score implies little involvement with environmental problems. High scoring people on the NEP Scale believe in fragility of nature’s balance, the possibility of an eco-crisis, and the limits of the earth to grow.

(Dunlap et al., 2000). High environmental concern implies that people are aware of various environmental problems and the consequences of their behavior, and they care about solving these problems (Minton and Rose, 1997). This shows that high environmental concern implies a positive attitude towards improving the environment. Ellen, Wiener, and Cobb-Walgren (1991) showed that a positive attitude towards the environment resulted in environmental friendly behavior such as recycling, attending public hearings, joining environmental groups and purchasing environmentally safe products. Furthermore, Ellen and colleagues (1991) found that people with an environmental friendly attitude were willing to reallocate their time, money and attention in order to improve their behavior towards a level which is better for the environment. This suits Azjen's and Fishbein's "correspondence rule" (1980), which implies high correlation between attitude and behavior when both are in line with each other. As a result, we assume that participants who score a high NEP value will rank environmental friendly products higher as compared to environmental unfriendly products. This also implies that we assume that low NEP value participants will rank environmental unfriendly products higher as compared to environmental friendly products, since this is in line with their attitudes.

Interestingly, literature shows the "low-cost hypothesis" (Diekmann and Preisendörfer, 2003), which states that environmental concern only influences ecological behavior under conditions which involve low-costs decisions which have little consequences for individuals. Diekmann and Preisendörfer (2003) showed that the effect on behavior of environmental concern disappears when the costs are high. Zintl (1989) states that low-cost situations should be distinguished from high-cost situations. In his paper, Zintl (1989) shows that high-cost situations evoke rational choice processes, while low-cost situations rely more on peripheral cues. Since participants in current research buy high-cost products (fridges), we assume that they will evaluate the products in a rational way. We assume that labels with a reference point will make more sense to the participants with a high NEP value. As a result, we assume that participants with a high NEP value rank the traffic light label and the checkmark label (labels with a reference point) higher as compared to labels with no reference point, which are the black and the colored score labels.

Finally, it has been shown that attention towards labels depends on consumer goals. Drichoutis and colleagues (2006) state that consumers who have health goals are more likely to use nutrition labels. If we generalize this finding, we expect that attention

towards carbon labels in this research is higher for consumers who are highly involved in environmental concerns. Based on this involvement, they will compare labels extensively and we assume that the positive or negative score of the label will be guiding and thus that brand versus product level has little influence on high NEP value consumers. However, as shown before, brands can help consumers in guiding purchase behaviour. Low NEP value participants will have little interest in comparing carbon labels, and thus we assume that they will use brand cues in their purchase behavior. As a result, we assume that low NEP value will give higher ratings of liking and attractiveness for labels on a brand level, compared to labels on a product level. This leads us to the following hypotheses:

H9: NEP value will influence participants' attitude towards the products in a way that high NEP value participants will prefer environmental friendly products, where low NEP participants will make no distinction between environmental friendly and unfriendly products.

H10: NEP value will influence participants' preference of labels in a way that low NEP value participants will prefer labels with a reference point, while there will be no difference in preference for the labels among high NEP participants.

H11: Low NEP value participants will prefer a label on a brand level over a label on a product label, while high NEP value participants will make no distinction between a brand versus a product level, but between positivity and negativity of the label.

Social Norms

Even though NEP Scale is worldwide the most widely used measure of environmental values, analysis of pro-environmental behavior showed that personal norms were far more correlated to pro-environmental behavior than NEP scale scores (Wiidegren, 1998). NEP and personal norms differ from each other in a way that NEP measures awareness of consequences, while personal norms intend to measure feelings of guilt related to harming the environment (Wiidegren, 1998). For this reason, we also want to investigate the moderating effects of norms. Minton and Rose (1997) found that both environmental attitude and personal norms evoked different processes among participants in their study. Results showed that personal norm had primary influence on

environmentally friendly behavior, while environmental attitude primary influenced intentions of behavior (Minton and Rose, 1997). Personal norms are tied to one's self-concept and refer to what people feel morally obligated to do (Cialdini, Reno, and Kallgren, 1990; Schwartz, 1973). Personal norms guide behavior since people have the desire to act in line with their own values. Hopper and Nielsen (1991) argued that personal norms may develop from existing societal view of how people ought to behave. Following this line of reasoning, it can be stated that personal environmental norms are internalized social norms (Minton and Rose, 1997). Cialdini and colleagues (1990) also placed personal norms, together with descriptive norms, under social norms. Cialdini et al. (1990) state that social norms are "*expectations of how people are supposed to act, think or feel in specific situations*". Descriptive norms describe the behavior of the majority and they state what kind of behavior can be considered typical (Cialdini et al., 1990). An example here can be "More than 80% of Dutch citizens buy carbon friendly products". However, we will adhere to the broader definition of social norms from Burchell, Rettie and Partel (2013), who state that norms can be seen as rules that guide and regulate social behavior in certain situations and that social norms indicate what is done by others and what we think that others do and approve (see also Cialdini et al., 1990). Social norms are believed to work because they give people an example to follow, whilst it would take more effort to set a leading example (Lewis and Neighbors, 2006). We argue that people who "follow" others in their behaviour elaborate less options and thus act less rational compared to people who set and rely on their own norms. As a result, we assume that people who score high on social norms are less rational than people who score low on social norms. This is in line with findings of Jacobsen, Mortensen and Cialdini (2011), who found that injunctive norms are most influential under higher levels of cognitive activity, whereas descriptive norms worked most effectively under lower conditions of cognitive activity. Based hereon, we assume that participants who score high on social norms will have a preference for labels with no reference point, while participants who score low on social norms will be more rational and thus evaluate the label via a more central route. This implies that participants who score low on social norms will prefer labels with a reference point and thus like the checkmark label and the traffic light label more as compared to the black- and colored score labels.

Continuing this line of reasoning, we also assume that less rational participants (the ones who score high on social norms) will rely on brands in order to increase their ease of choice. More rational participants, scoring low on social norms, will evaluate the aspects

of labels on a brand versus on a product level in a similar way. This leads to the following hypotheses:

H12: Beliefs in social norms will influence participants' attitude towards the products in a way that participants who score high on social norms will prefer products with a label without a reference point, while participants who score low on social norms will prefer products with a label with a reference point.

H13: High beliefs in social norms will influence participants' ratings of liking and attractiveness towards brand versus product level in a way that they will prefer a label on a brand level. Low scoring participants will evaluate the labels on a brand versus on a product level the same in terms of liking, attractiveness and willingness to pay.

Perceptions of influence

Perceptions of influence is a concept which refers to the extent to which people believe that their behavior can influence a certain occurrence. Research has shown that employees' perceptions of influence, and thus expectations with regard to participation in decision making, can lead to increased satisfaction, trust, achievement, commitment, and involvement (Hrebiniak, 1974). This shows that perceptions of influence can change behavior. To our knowledge, perceptions of influence is not elaborated yet in combination with environmental concern. This is odd, since Dolan and McGeorge (1994) found that high identification with athletic competitions and sports clubs could influence spectators' perceptions of influencing the result of sport events. In a general sense, this shows that involvement can lead to the idea among people that they can influence a result. This can be really interesting, since people who care about the environment are in general highly involved in terms of consumption and behavior. In current research, this involvement might lead to high perceptions of influence, which implies that participants feel that they can do something in order to improve the quality of the environment. It can be seen that having high perceptions of influence leads to emotional states such as increased feelings of trust, commitment and self-control (Patchen, 1970). Dolan and McGeorge (1994) also showed that highly involved sport fans reported an increase in pre- and postgame positive emotions, as a result of a victory. On the contrary, low involved spectators showed minimal emotional changes. Drawing forward on these findings, we assume that results in line with participants' values lead to increased positive emotions, and thus that participants with high perceptions of influence will be

more satisfied with environmental friendly products containing labels without a reference point (black- and colored score labels) as compared to environmental friendly products containing labels with a reference point (the checkmark and traffic light labels). Based on this assumption that participants with high perceptions of influence are more satisfied, we also assume that labels without a reference point will be more meaningful to them, and thus that they are willing to pay more for products with these labels.

Furthermore, Algesheimer and colleagues (2005) showed that the increased positive emotions lead to increased identification with a brand. As a result, we assume that participants with high perceptions of influence will rate labels on a brand level higher as compared to labels on a product level. We assume that this applies for positive as well as for negative labels. Low perceptions of influence will result in more rational behavior and thus we expect that participants with low perceptions of influence will evaluate the labels more rational, with special attention towards the score of the label. As a result, we expect no differences in preferences for brand versus product level for low scoring participants on perceptions of influence. This leads us to the following hypotheses:

H14: Perceptions of influence will direct participants' attitude towards products in a way that participants who score high on perceptions of influence will prefer products with a label without a reference point, while participants who score low on perceptions of influence will prefer products with a label with a reference point.

H15: Perceptions of influence will direct participants' willingness to pay for products in a way that participants who score high on perceptions of influence will be willing to pay more for labels without a reference point as compared to participants who score low on perceptions of influence.

H16: High perceptions of influence will lead to increased preferences for labels on a brand level as compared to labels on a product level. Participants with low perceptions of influence will show no difference in preferences for brand versus product level.

Method

Participants

For this Master thesis, an online survey was performed with the programme Qualtrics. With reaching as many respondents as possible in mind, the survey was created in English. Another reason for choosing an English survey was the fact that the literature regarding the used constructs was mainly available in English. Translation could induce misinterpretation or other errors of the questions. Since the target group were mainly students at the Wageningen University, where classes are generally taught in English, language should not be any problem. The survey has been distributed by means of e-mail and online social media. It was tried to attract new respondents based on networks of respondents who already wanted to fill in the survey. Friends and acquaintances were asked to share the survey with their friends and acquaintances, and so on. The link to the survey was also shared in different Facebook groups. Since two iPad's were distributed among the participants, people were eager to fill in the survey.

Regarding respondents, it was aimed to reach at least 30 participants per between subjects condition, which should sum up to a total of 120 respondents. A total of 210 respondents was reached, which mainly were students from Dutch descent. After exclusion of incomplete ($N = 52$) and incorrect responses ($N = 27$), a total of 131 respondents remained. Ages varied between 18 and 37 years, with a mean of 22 years ($SD = 2.76$).

Design

In this study, a 4(type of label: checkmark versus traffic light versus black score versus colored score) x 2(label level: brand versus product) x 2(type of product: utilitarian versus hedonic) mixed model design was used. However, current study only focuses on the part where type of label (4) and NEP value are the between subjects variables, and on label level (2, brand versus product), which are the within subjects variables. We also analysed the effects of perceptions of influence and social norms, as controlling variables. Qualtrics randomly but evenly assigned all the survey participants to one of the four types of labels.

Procedure

The experiment began with an introduction and informed consent. Hereafter, participants had to fill in questions related to their New Ecological Paradigm score (NEP), perceptions of climate change, beliefs in global climate change, social and personal norms, and need for cognition. Then, participants in the survey were randomly exposed to one of the four types of carbon labels. This could be a checkmark, a traffic light, a black score or a colored score. Every respondent saw one of the four labels four times. The first time, the label was on a brand level, positively rated. The second time, the label was also on a brand level, but negatively rated. The third time, the label was on a product level, positively rated. The fourth time, the label was also on a product level, but negatively rated (see appendix I for examples).

Labels

Regarding the checkmark label, positive ratings imply a green checkmark, while negative ratings were a red x-sign. For the traffic light label, positive ratings were a green traffic light, while negative ratings were a red traffic light. Regarding the black score, positive ratings were high scores under the hundred. In this research, scores of 87 (brand level) and 85 (product level) were used. Negative ratings for a black score were the scores of 29 (brand) and 27 (product). The same ratings were used for the colored score label. However, in this case were the high ratings (87 and 85) displayed in a green color, while the negative ratings (29 and 27) were displayed in a red color (appendix II).

Intro text

All the conditions (type of label) began with a short introductive text, which said *“Recently you have decided that your kitchen needs remodelling. After spending some time making decisions about all of the appliances, the only item that remains to be chosen is a refrigerator. After several hours of searching and comparing you end up with 2 different refrigerators within your budget”*. After this text, which was supposed to put the participants in a comfort situation in which they had to make a decision, a text was shown in order to emphasize the fact that the first label they would see contained information about carbon emissions of the fridge on a brand level. This text stated: *“One of the features is a carbon label, which indicates the carbon emission friendliness. Notice that the carbon label is on a brand level, which indicates the score of the brand for all of its products combined. This score does not necessarily apply to this specific refrigerator.*

The green checkmark indicates that it is a carbon friendly brand, while the red checkmark indicates that it is a carbon unfriendly brand”.

Manipulation checks

The next question after this intro text was a manipulation check to see whether the respondents had realized what they were going to buy and that the information on the label was telling something about the brand of the fridge as a whole, and not about the specific fridge. Participants were asked “what will you buy in the next question” and “What is the carbon label about?”. The two optional answers on the latter question were: *“The carbon label provides information about the carbon emission friendliness of this specific product”* and *“The carbon label provides information about the carbon emission friendliness of the brand, which applies to all products of this brand combined”*. If respondents incorrectly answered one of these question, their data was deleted from the dataset.

Dependent variables: brand level

After the manipulation check, the respondents saw the first picture of a fridge with a positive carbon rating on a brand level. Participants were told that *“Prices of these fridges vary between €400 and €600”* and they were asked how much they were willing to pay for the fridge. Hereafter, they saw a smiley face (5-point scale) and were asked to indicate how much they liked the fridge. Finally, they were asked how attractive they considered the fridge on a 7-point Likert scale. For pictures of the fridges with the four different types of labels with a positive carbon rating on a brand level, see appendix III.

After the respondents answered the questions for the positive carbon rating on a brand level, the same questions were asked for a fridge with a negative carbon rating on a brand level (see appendix IV for an example).

Dependent variables: product level

After these questions, which compared positive versus negative carbon ratings on a certain type of carbon label (checkmark, traffic light, black score or colored score) on a brand level, a new transcript was displayed. This text intended to show the respondents that the next questions would be about a fridge with a carbon label on a product level, which showed information about the specific fridge. This text stated: *“One of the features*

is a carbon label, which indicates the carbon emission friendliness. Notice that the carbon label is on a product level, which indicates the score on a product level. This score does not per se apply for other similar products. The green checkmark indicates that it is a carbon friendly product, while the red checkmark indicates that it is a carbon unfriendly product". After this text, a new manipulation check was conducted. Again, incorrect answers resulted in deletion of the participants from the dataset. The first picture after this text showed a fridge with a positive carbon rating on a product level. Participants were told that "*Prices of these fridges vary between €400 and €600*" and they were asked how much they were willing to pay for the fridge. Hereafter, they saw a smiley face (5-point scale) and were asked to indicate how much they liked the fridge. Finally, they were asked how attractive they considered the fridge on a 7-point Likert scale. For pictures of the fridges with the four different types of labels with a positive carbon rating on a product level, see appendix V.

After the respondents answered the questions for the positive carbon rating on a product level, the same questions were asked for a fridge with a negative carbon rating on a product level (see appendix VI for an example).

Label assessment

After answering the questions regarding attractiveness, liking and willingness to pay, the main part of the survey was completed. Next were some questions about the type of label which the respondents saw (appendix VII). These 7-point Likert scale questions were asked in order to assess participant's opinion towards the different labels. This construct contained 8 questions which were high reliable together ($\alpha = .753$). Reliability analysis showed that deleting the item "the information that the carbon label provides is difficult to interpret" would make the Cronbach's Alpha to .825 (table 1). If this scale was recoded and thus the scores were reversed instead of deleting the item as a whole, Cronbach's Alpha increased from .753 to .779. In this case, two times four items remained. The first four items measured ease of interpretation of the different labels, while the last four items measured the influence on decisions of the different types of labels. Descriptives of the results of the questions can be found in appendix XIII.

Table 1. Reliability Item-Total Statistics about label assessment

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item- Total Correlation	Cronbach's Alpha if Item Deleted
The carbon label gives important information	33,15	36,684	,703	,723
The carbon label makes me more aware of the aim of reducing carbon emissions	33,02	37,623	,573	,741
The carbon label influences my decision	33,14	36,304	,552	,742
The information that the carbon label provides is difficult to interpret	34,02	44,146	,077	,825
I take the carbon label seriously	33,36	36,493	,599	,735
I understand the information on the carbon label	32,89	38,548	,526	,748
The carbon label provided me with enough information to estimate the carbon emission friendliness of the brand	34,19	37,986	,450	,760
The carbon label provided me with enough information to estimate the carbon emission friendliness of the product	33,56	36,694	,511	,749

Recall

Label assessment questions were followed by questions about recall of the carbon ratings on the different types of labels. Participants in the checkmark condition saw the fridge of the first displayed picture (with a positive carbon rating on a brand level), but in this picture was the place of the carbon label empty. The question here was “*Please look at the picture below, this was the first product that you evaluated. Please indicate the*

label that was given to this product". This was done for all the fridges which they saw in the questions before (and thus four times; for the positive and negative ratings on a brand level and on a product level, see appendix VIII). Respondents in the traffic light label condition saw the same question, but with different answer alternatives (appendix IX). For the black score label, respondents were asked: *"Please look at the picture below, this was the first product that you evaluated. Please indicate the score of the label that was given to this product, if you do not know the exact score, please type in your best estimate"*. See appendix IX for an example. Regarding the colored score, participants were asked the same questions as for the black score label, but now they were also asked what color the label had (appendix X). These questions measured the exact recall. Since it was expected that it would be really difficult to remember an exact number, compared to a "2"-point scale (checkmark) or a "3"-point scale (traffic light), a second test will be performed. In this test, participants' answers for the score labels (black and colored) were recoded. An answer on negative carbon labels was recoded into correct (a zero) if the answer was between 0 and 33. An answer on positive carbon labels was recoded into correct (a zero) if the answer was between 67 and 100. All other answers were coded into incorrect (a one). This recoded score is thus an indication of "approximately right" scores and will be called "recall of category".

Debriefing

The experiment ended with a number of socio-demographic variables (e.g. gender, age and study) and a debriefing. In this part, a short explanation about the purpose of the study and confidentiality was given (appendix XI).

Results

The following chapter contains the results of current research, starting with correlation checks and scale reliability tests. Hereafter, the analysis of the different types of carbon labels will be covered. Four types of labels (checkmark, traffic light, black score, colored score) were compared. The labels varied in rating (positive versus negative) and level of information (brand versus product level). Multiple operationalizations, in particular WTP, liking and attractiveness are used in order to measure the hypotheses.

Analysis

Mixed model ANOVA analysis were used to determine group differences between willingness to pay (WTP hereafter), liking, and attractiveness. Levene's Test of Equality of Error Variances always gave a non-significant result, which implies that indicated that variances between groups were homogeneous, and so the analysis may be conducted (Field, 2009).

Since two conditions at the time were compared (carbon label on a brand level versus on a product level), the Mauchly's Test of Sphericity was not able to show a significance level. Field (2009) states that sphericity is met if the repeated-measures variable has only two levels, which is the case for this analysis. This applies for all tests analysing the labels with information on a brand versus a product level, and so we will look at the uncorrected *F*-ratios.

Covariates and correlation

In order to select the covariates to work with, a correlation test was performed (table 2). Harlow (2014) states that there should be low correlations among covariates in order to avoid collinearity problems like instability problems and biases. This implies that covariates should be relatively uncorrelated with each other. Pearson correlation shows that a small relation has a *r*-value smaller than .30. A medium correlation lies between .30 and .50, and a large correlation has a higher *r*-value than .50 (Lund Research Ltd, 2013).

Table 2. Correlation matrix of covariates

	NEP	Perceptions of influence	Perceptions of responsibility	Beliefs about climate change	Personal norms	Social norms	Need for cognition
NEP							
Pearson correlation	1	.406**	.551**	.638**	.653**	-.038	.483**
Sig. (2-tailed)		.000	.000	.000	.000	.664	.000
Perceptions of influence							
Pearson correlation	.406**	1	.377**	.284**	.374**	.208*	.363**
Sig. (2-tailed)	.000		.000	.001	.000	.017	.000
Perceptions of responsibility							
Pearson correlation	.551**	.377**	1	.413**	.444**	.084	.321**
Sig. (2-tailed)	.000	.000		.000	.000	.343	.000
Beliefs about climate change							
Pearson correlation	.638**	.284**	.413**	1	.574**	-.009	.508**
Sig. (2-tailed)	.000	.001	.000		.000	.920	.000
Personal norms							
Pearson correlation	.653**	.374**	.444**	.574**	1	-.223*	.524**
Sig. (2-tailed)	.000	.000	.000	.000		.010	.000
Social norms							
Pearson correlation	-.038	.208*	.084	-.009	-.223*	1	-.003
Sig. (2-tailed)	.664	.017	.343	.920	.010		.972
Need for cognition							
Pearson correlation	.483**	.363**	.321**	.508**	.524**	-.003	1
Sig. (2-tailed)	.000	.000	.000	.000	.000	.972	

** . Correlation is significant at the .01 level (2-tailed).

* . Correlation is significant at the .05 level (2-tailed).

Since current research has a small sample size, it has been decided not to add too many covariates. NEP appeared to have a positive correlation with all the other covariates, except for social norms. These correlations were all medium to large ($r > .30$, Lund Research Ltd, 2013). Perceptions of influence had a positive correlation with all of the

covariates, all being medium, except for social norms and beliefs about climate change, which were small ($r < .30$).

Perceptions of responsibility had a medium relationship ($r > .30$) with all of the covariates as well, except for social norms. This correlation was small ($r < .30$).

Beliefs of climate change also had an significant positive medium to large relation with all the covariates, except for social norms. The correlation with social norms was negative and small ($r < .30$).

Personal norms correlates large ($r > .50$) to NEP, beliefs of climate change and need for cognition. It correlated medium ($r < .50$) with perceptions of influence and perceptions of responsibility, and it had an negative, small correlation with social norms ($r < .30$).

Need for cognition had a medium to high correlation with all the other covariates ($r > .30$), except for social norms ($r < .30$).

It becomes clear that social norms had a small or no significant correlation with most of the other covariates. This implies that social norms can be seen as an independent covariate which does not correlate with the other ones. This is interesting since we are interested in using low correlated covariates, in order to overcome collinearity problems (Harlow, 2014). Since collinearity problems would imply that estimates of multiple regression will largely change as a result of small changes in the model, which reduces the predictive power of the model (Næs and Mevik, 2001), it has been decided to work with the low correlating covariates NEP, perceptions of influence and social norms. In order to run tests with these covariates, it was decided to make single constructs of the variables. After checking the reliability of the scales (table 3), all the covariates were merged based on the median split method. As a result, a distinction could be made between high and low NEP participants, high and low feelings of perceptions of influence, and high and low social norms.

Table 3. Reliability of covariates

	Number of items	Cronbach's Alpha	Mean	S.D.
NEP	15	.875	4.402	.887
Perceptions of influence	3	.844	4.262	1.294
Social Norms	4	.808	3.687	1.161

Statistical analysis

The test of Repeated Measures has been chosen as the test that provides the most comprehensible and analytically sound output to interpret. Condition (type of label) and NEP value were always the between-subjects factors, and social norms and perceptions of influence were the controlling variables. In the analysis, mean scores of the controlling variables social norms and perceptions of influence were used. This was done in order to provide more detailed results, because data was lost when the mean scores were divided into only two groups, causing the variance to drop dramatically. As a result, using the mean scores for the controlling variables creates a better picture and a more sound effect. The between subjects factor NEP value was recoded into high versus low scores based on a median split. This was done in order to run a repeated measures analysis with a meaningful post hoc output. If the mean score of NEP was used, too many levels would have been created, which would end up in uninterpretable results. Two levels (high versus low) appeared to give strong results.

In this results section, within subjects effects will be covered at first. If there appears to be any effect for the controlling variables, the direction and p -value of this correlation will be checked by means of a linear regression test. In this regression test, we always inserted both the covariates social norms and perceptions of influence, in order to reduce noise. Hereafter, the between subjects effects will be evaluated. It is important to note that the between subjects tests are based on average scores of the within subject variables. As such, these tests yield no information about the within-subjects effects (brand versus product level, or positive versus negative ratings). However, these differences can be found in the results of the post hoc test. This is why the results of both the within subjects and between subjects tests will be discussed.

An important concluding remark is that due to the fact that this study is explorative and not confirmative, the alpha criterion is raised from .05 to .10. This was done in order to find more statistical significances and provide more levels that future research can further explore.

WTP analysis

Next section will evaluate the analysis of participant's WTP for the four different types of carbon labels. A comparison will be made between positive versus negative ratings of the label and between the label being on a brand level versus on a product level.

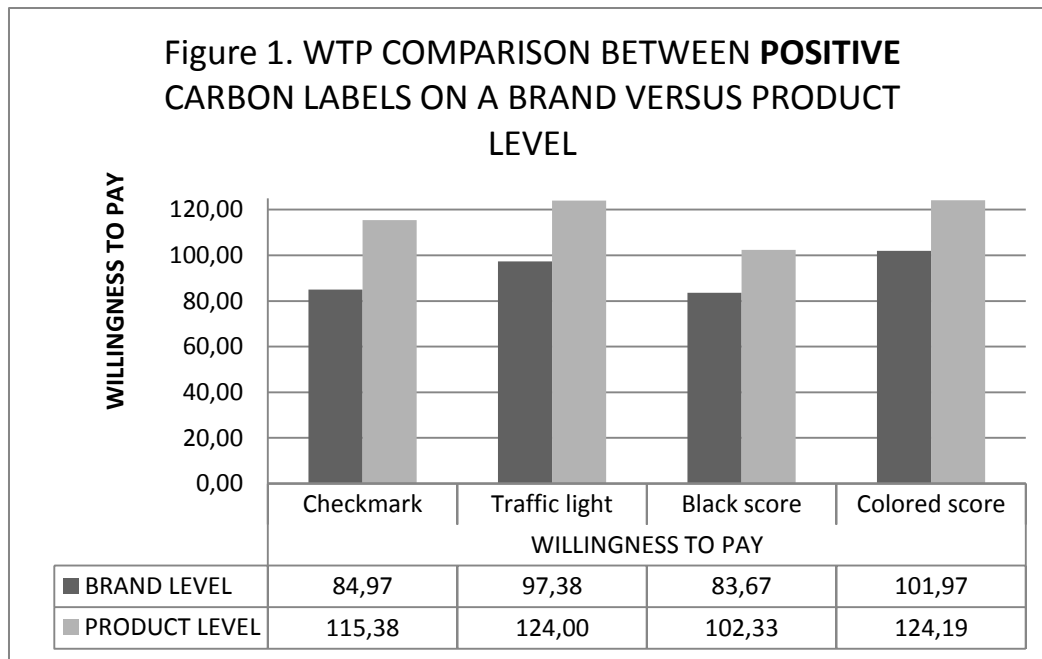
WTP Brand versus Product level, positive labels

We ran an analysis with participant's WTP for fridges with a positive carbon label on a brand and on a product level as dependent variable. Condition and NEP were the predictors, and perceptions of influence and social norms were taken as covariates. This analysis reveals the following:

Within subjects

The Tests of Within-Subjects Effects shows a significant main effect ($F(1, 121) = 8.432, p = .004$) of the information on the positive carbon label on a brand level ($M = 92.02, SD = 49.79$) versus on a product level ($M = 116.48, SD = 50.22$). This effect tells us that, if we ignore the fact whether the rating was performed for four different types of labels, the WTP for fridges with a carbon label on a brand level differed significantly from the WTP for fridges with a carbon label on a product level. It can be seen that people are willing to pay more for a fridge with a positive carbon label on a product level as compared to a fridge with a positive carbon label on a brand level. This is not in line with hypothesis 6.

There is no significant main effect of condition on WTP ($F(3,121) = 1.200, p = .313$). This implies that participants were not willing to pay significantly more for certain carbon labels on a brand versus on a product level. Yet is interesting to see that there are quite some differences between the four types of label. Means of WTP for the four different positive carbon labels on a brand versus product level are displayed in figure 1.



Furthermore, NEP value had a significant main effect on WTP for fridges with a carbon label on a brand level versus a carbon label on a product level, $F(1,121) = 10.125$, $p = .002$. This shows that the ratings for WTP between respondents with a high or low NEP significantly differed from each other. Participants with a low NEP were rating their WTP for a fridge with a carbon label on a brand level on average less ($M = 82.61$, $SD = 53.45$) compared to participants with a high NEP ($M = 101.57$, $SD = 44.16$). Also, low NEP participants were willing to pay on average less ($M = 98.91$, $SD = 54.39$) as compared to participants with a high NEP ($M = 134.32$, $SD = 38.43$) for a fridge with a carbon label on a product level. This shows that both high and low NEP participants prefer labels on a product level. This is in not in line with hypothesis 11.

No significant effects of the controlling covariates perceptions of influence ($F(1,121) = 1.862$, $p = .175$) and social norms ($F(1,121) = .003$, $p = .956$) were found. This implies no effect of these variables on consumers' preferences, which is in contrast with hypotheses 12 and 15.

Between subjects

Regarding the four different labels, no significant simple main effect of condition was found, $F(3,121) = 2.033$, $p = .113$. This effect tells us that participants did not significantly differed in their indicated WTP for fridges with whether a checkmark, traffic light, black score, or colored score label. This finding is contradicting to hypotheses 3 and 10.

NEP had a significant simple main effect between subjects $F(1,121) = 10.714, p = .001$. This effect tells us that participants' NEP had an effect on their WTP. Bonferroni corrected post hoc tests showed that ratings for WTP between respondents who scored high ($M = 117.28$) versus low ($M = 90.66$) on NEP did significantly differ ($p = .001$). Respondents with high NEP score wanted to pay on average €26.62 more for a fridge with a positive carbon label as compared to respondents who scored low on NEP. This is in contrast with hypothesis 9.

No significant effects of the controlling covariates perceptions of influence ($F(1, 121) = 1.365, p = .245$) and social norms ($F(1,121) = .100, p = .752$) were found.

Finally, a simple main effect of information on positive carbon labels on a brand level versus on a product level was found. A significant difference could be found between the WTP for a fridge with a positive carbon label on a brand level ($M = 91.67, SE = 4.30$) versus a fridge with a positive carbon label on a product level ($M = 116.27, SE = 4.14$). So, on average, participants were willing to pay €24.60 euro's more for a fridge with a positive carbon label on a product level, as compared to a fridge with a positive carbon label on a brand level ($p = .000$). As opposed to the within-subjects means and standard errors which are mentioned before, latter mentioned numbers are corrected for perceptions of influence and social norms. This is, again, in contrast with hypothesis 6.

WTP Brand versus Product level, negative label

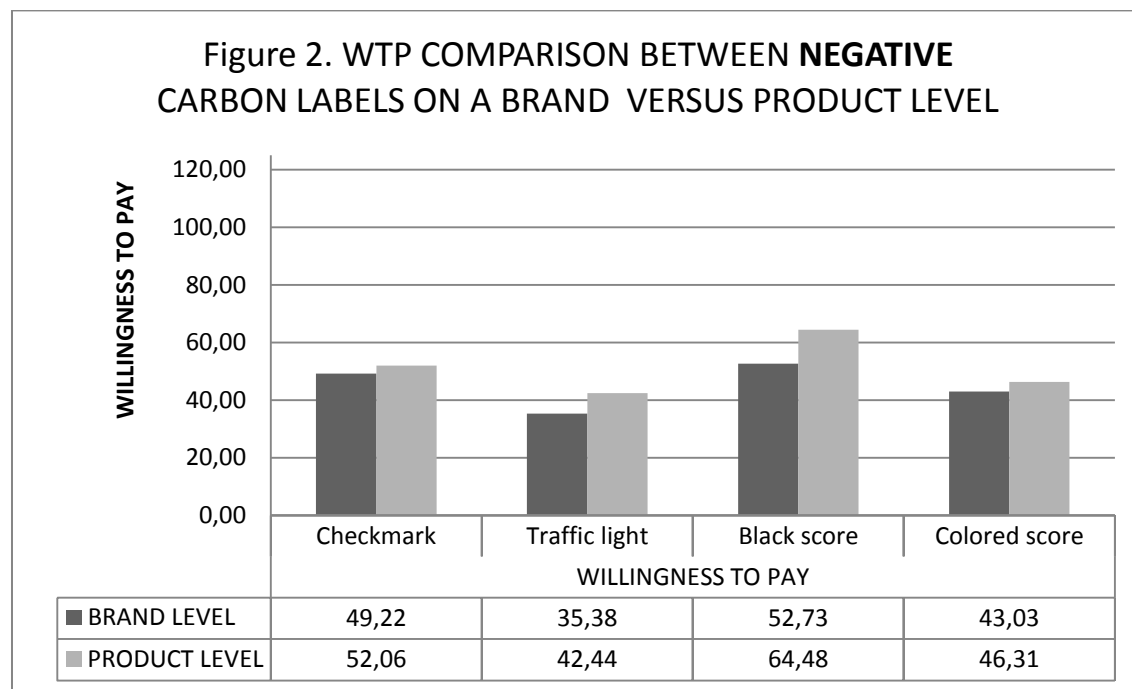
The same analysis was conducted for WTP, with negative carbon labels on a brand and on a product level as dependent variable. Condition and NEP were the predictors, and perceptions of influence and social norms were taken as covariates. This analysis reveals the following:

Within subjects

While the positive carbon labels on a brand versus a product level significantly differed from each other, no significant difference was found between the WTP for fridges with a negative carbon label on a brand versus a product level ($F(1,121) = .129, p = .720$). For a fridge with a negative carbon label on a brand level, participants were willing to pay €45 euro's ($SD = 47.31$) as compared to €51.29 ($SD = 46.45$) for a fridge with a negative carbon label on a product level. This not-significant effect tells us that there is no difference for WTP between fridges with a negative carbon label on a brand versus a

product level. This is in contrast with hypothesis 7. This might imply that participants do not see a difference between brand versus product, when the label is negative.

Condition had no main effect ($F(3,121) = .634, p = .595$) on the WTP ratings between brand versus product level. Mean differences in WTP between the four different negative carbon labels on a brand versus product level are displayed below, in figure 2.



Furthermore, no significant effects within subjects were found. Perceptions of influence had no effect as a controlling covariate ($F(1,121) = .002, p = .968$) and no significant effect of the controlling covariate social norms ($F(1,121) = .009, p = .924$) was found. No significant main effect of NEP on WTP was found ($F(1,121) = 2.675, p = .105$).

Between subjects

The between subjects table does not show any significant results. Condition ($F(3,121) = .930, p = .428$), NEP ($F(1,121) = .343, p = .559$) as predicting variables, and perceptions of influence ($F(1,121) = 1.532, p = .218$) and social norms ($F(1,121) = .811, p = .370$) as controlling covariates did not have any effect. This shows that, in line with the findings among positive carbon labels, label type (checkmark, traffic light, black score or colored score) does not influence participant's WTP when the carbon information is a negative indication ($F(3,121) = .930, p = .428$).

Interestingly, if the means for WTP are corrected for perceptions of influence and social norms, the difference between these average WTP for fridges with a negative label on brand versus a product level becomes significant and thus a simple main effect of negative carbon labels on a brand versus on a product level was found. The WTP for a fridge with a negative carbon label on a brand level ($M = 44.13$, $SE = 4.17$) differed significantly from the WTP for a fridge with a negative carbon label on a product level ($M = 50.48$, $SE = 4.09$). This shows that, on average, participants were willing to pay €6.35 euro's more for a fridge with a negative carbon label on a product level, as compared to a fridge with a negative carbon label on a brand level ($p = .046$). This supports hypothesis 7. However, we need to take a more intensively look to our results, before we can accept or reject this hypothesis.

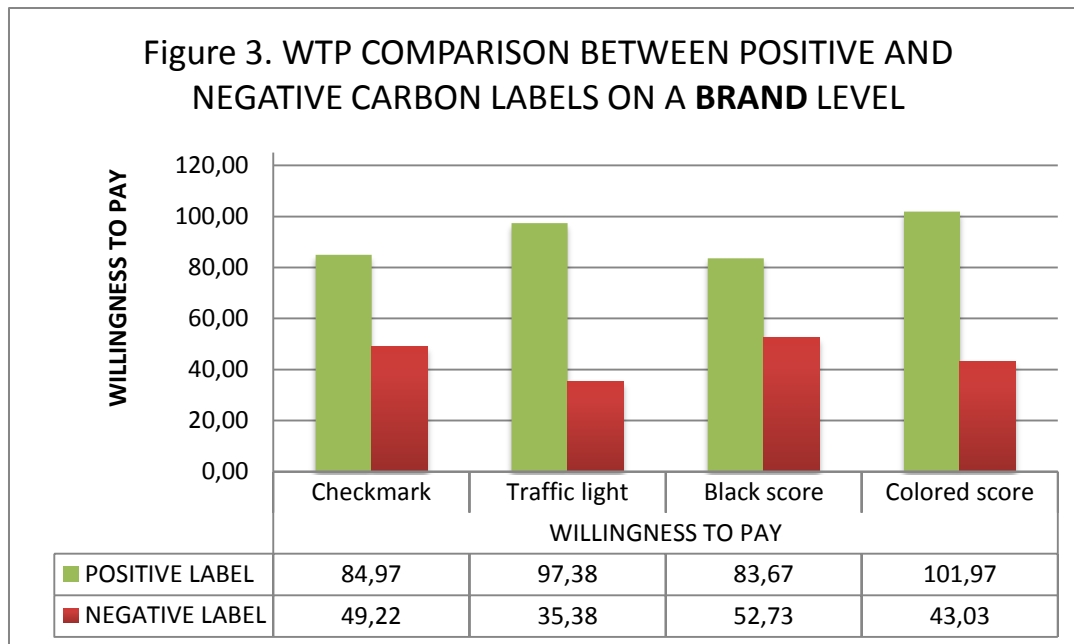
WTP Brand level, positive versus negative labels

The participant's willingness to pay for fridges with positive and negative carbon labels on a brand level was analysed with condition and NEP as predictors, and perceptions of influence and social norms as covariates, the results reveal the following:

Within subjects

No significant difference was found between the WTP for fridges with whether a positive or a negative carbon label on a brand level ($F(1,121) = .605$, $p = .438$). For a fridge with a positive label on a brand level, participants were willing to pay on average €92.02 ($SD = 49.79$) as compared to €45.00 ($SD = 47.31$) for a fridge with a negative label on a brand level. This tells us that there is no significant difference between the WTP for fridges with whether a positive or a negative carbon label on a brand level.

A significant main effect of condition was found ($F(3,121) = 3.837$, $p = .012$). This implies that condition has an effect on the difference between WTP for a fridge with a positive versus a negative carbon label on a brand level. Mean differences in WTP between the positive and negative carbon labels on a brand level are displayed in figure 3.



The controlling variable perceptions of influence correlated with WTP, $F(1,121) = 7.925$, $p = .006$. This influence is in a positive direction for the positive carbon label ($B = .207$, $p = .021$), while it is in a negative direction for a negative carbon label ($B = -.053$, $p = .558$).

Social norms had no effect as a controlling covariate ($F(1,121) = .274$, $p = .602$). NEP had no main effect ($F(1,121) = .593$, $p = .443$) on the WTP.

Between subjects

Condition had no significant effect on participant's WTP for fridges with carbon labels on a brand level ($F(3,121) = .219$, $p = .883$). This implies that the willingness to pay between the four labels (checkmark, traffic light, black score, and colored score) was about the same.

NEP appeared to have a significant simple main effect on the WTP for fridges with whether a positive versus negative carbon label on a brand level ($F(1,121) = 2.958$, $p = .088$). Bonferroni corrected post hoc tests showed that ratings of WTP between respondents who scored high ($M = 74.51$, $SE = 5.35$) versus low ($M = 61.29$, $SE = 5.33$) on NEP did significantly differ. Respondents with a high NEP score were willing to pay on average €13.22 more for a fridge as compared to respondents who scored low on NEP.

Perceptions of influence ($F(1,121) = .065, p = .799$) and social norms ($F(1,121) = .359, p = .550$) as controlling covariates did not have any effect on the WTP.

Interestingly, if the means for WTP are corrected for perceptions of influence and social norms, the difference between the WTP for fridges with whether a positive or a negative carbon label becomes significant. This shows a simple main effect of a positive versus a negative carbon label on a brand level. A significant difference could be found between the WTP for a fridge with a positive carbon label on a brand level ($M = 91.67, SE = 4.30$) versus a fridge with a negative label on a brand level ($M = 44.13, SE = 4.17$). This shows that, on average, participants were willing to pay €47.54 euro's more for a fridge with a positive carbon label as compared to a fridge with a negative carbon label on a brand level ($p = .000$).

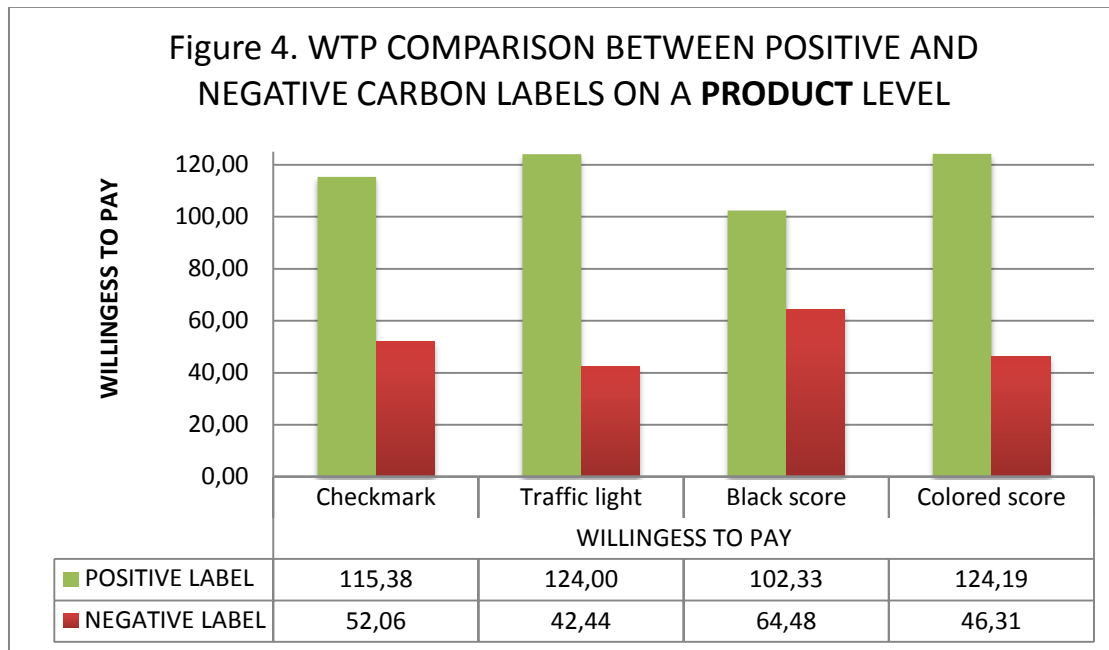
WTP Product level, positive versus negative labels

Additionally, an analysis with WTP for fridges with positive and negative carbon labels on a product level was conducted. In this case, condition and NEP were still predictors, and perceptions of influence and social norms covariates. This analysis reveals the following:

Within subjects

The Tests of Within-Subjects Effects shows a significant main effect of a positive carbon label ($M = 116.48, SD = 50.22$) versus a negative carbon label ($M = 51.29, SD = 46.45$) on a product level $F(1, 121) = 7.296, p = .008$. This effect tells us that, if we ignore the fact whether the rating was performed for different types of labels, the ratings of WTP for fridges with a positive carbon label on a product level still significantly differed from ratings of WTP for fridges with a negative carbon label on a product level. It can be seen that people are willing to pay more for a fridge with a positive carbon label on a product level as compared to a fridge with a negative carbon label on a product level.

A significant main effect of condition was found ($F(3,121) = 6.824, p = .000$). This implies that condition has an effect on the difference between WTP for a fridge with a positive versus a negative carbon label on a product level. This means that at least on pair out of the four labels (checkmark, traffic light, black score, and colored score) significantly differed from each other regarding the indicated willingness to pay. Mean differences in WTP between the four different positive and negative carbon labels on a product level are displayed in figure 4.



Furthermore, NEP value had a significant main effect on WTP for positive versus negative carbon labels on a product level, $F(1,121) = 19.779$, $p = .000$. This shows that the willingness to pay between respondents with a high or low NEP significantly differed from each other. Participants with a low NEP ($M = 98.91$, $SD = 54.39$) versus a high NEP ($M = 134.32$, $SD = 38.43$) indicated different WTP's for a fridge with a positive carbon label on a product level. Also for a fridge with a negative carbon label on a product level differed the WTP's between participants with a low NEP ($M = 52.10$, $SD = 47.75$) versus a high NEP ($M = 50.48$, $SD = 45.45$). This indicates that both high- and low NEP participants made a distinction between environmental friendly and unfriendly fridges. This is in contrast with hypothesis 9.

The controlling variable Perceptions of influence appeared to correlate with WTP, $F(1,121) = 3.092$, $p = .081$. This influence is in a positive direction for the positive carbon label on a product level ($B = .163$, $p = .070$), while it is in a negative direction for a negative carbon label on a product level ($B = -.091$, $p = .311$).

Social norms had no effect as a controlling covariate ($p = .579$) on the WTP.

Between subjects

Condition had no significant effect on participant's WTP for fridges with positive or negative carbon labels on a product level ($F(3,121) = .167$, $p = .918$) and thus the

willingness to pay between the four labels (checkmark, traffic light, black score, and colored score) was on average the same, after correction for perceptions of influence and social norms. This again implies that, when a label is on a brand level, little distinction is being made between a positive versus negative level.

As a result of the between subjects test, NEP appeared to have a significant simple main effect on the WTP for fridges with whether a positive versus negative carbon label on a product level ($F(1,121) = 5.922, p = .016$). Bonferroni corrected post hoc tests showed that ratings of WTP between respondents who scored high ($M = 92.40, SE = 5.16$) versus low ($M = 74.36, SE = 5.14$) on NEP did significantly differ. Respondents with a high NEP score rated their WTP for a fridge on average €18.04 higher as compared to respondents who scored low on NEP.

Perceptions of influence ($p = .761$) and social norms ($p = .493$) as controlling covariates did not have any effect on the ratings of WTP.

Finally, after correcting the mean WTP for perceptions of influence and social norms, a simple main effect of positive versus negative carbon labels on a product level was found. A significant difference could be found between the WTP for fridges with a positive carbon label on a product level ($M = 116.27, SE = 4.14$) versus a fridge with a negative carbon label on a product level ($M = 50.48, SE = 4.09$). On average, participants were willing to pay €65.80 euro's more for a fridge with a positive carbon label on a product level as compared to fridge with a negative carbon label on a product level ($p = .000$).

Conclusion WTP

Based on the results in the previous sections, we can state the following:

H3: Carbon emission labels containing a checkmark and a traffic light label will be preferred over a score label (black or colored).

No effect of type of label was found for neither positive nor negative carbon labels with regard to WTP. This implies that the type of label (check mark, traffic light, black score or colored score) does not affect participant's WTP, meaning that participants were willing to pay about the same amount of money among the different labels. As a result, **hypothesis 3** will be **rejected**.

H6: Respondents are willing to pay more for a positive label on a brand level compared to a positive label on a product level.

For both the positive and negative carbon labels, participants were willing to pay more for a fridge with a carbon label on a product level as compared to a fridge with a carbon label on a brand level. This is in contrast with **hypothesis 6**, which as a result will be **rejected**.

H7: Respondents are willing to pay more for a negative label on a product level compared to a negative label on a brand level.

Participants indicated not to have a different WTP for fridges with a negative carbon label on a brand versus a product level. This is in contrast with the assumption that was made in **hypothesis 7**, which will be **rejected**.

H9: NEP value will influence participants' attitude towards the products in a way that high NEP value participants will prefer environmental friendly products, where low NEP participants will make no distinction between environmental friendly and unfriendly products.

The willingness to pay between respondents with a high or low NEP differed significantly from each other. This shows that both high- and low NEP participants make a distinction

between environmental friendly and unfriendly fridges. This is in contrast with **hypothesis 9**, which as a result will be **rejected**.

H10: NEP value will influence participants' preference of labels in a way that low NEP value participants will prefer labels with a reference point (checkmark, traffic light), while there will be no difference in preference for the labels among high NEP participants.

Participants made no significant distinction in their indicated WTP for fridges with whether a checkmark, traffic light, black score, or colored score label. This finding is contradicting with **hypothesis 10**, which will be **rejected**.

H11: Low NEP value participants will prefer a label on a brand level over a label on a product label, while high NEP value participants will make no distinction between a brand versus a product level, but between positivity and negativity of the label.

Both high and low NEP participants prefer labels on a product level. This is in not in line with **hypothesis 11**, which thus will be **rejected**.

No significant effects of the controlling covariates perceptions of influence and social norms were found. This implies no effect of these variables on consumers' preferences, which is in contrast with **hypotheses 12 and 15**. Both of them will be **rejected**.

H12: Beliefs in social norms will influence participants' attitude towards the products in a way that participants who score high on social norms will prefer products with a label without a reference point, while participants who score low on social norms will prefer products with a label with a reference point.

H15: Perceptions of influence will direct participants' willingness to pay for products in a way that participants who score high on perceptions of influence will be willing to pay more for labels without a reference point as compared to participants who score low on perceptions of influence.

Taken this together, it appears that hypotheses 3, 6, 7, 9, 10, 11, 12, and 15 are rejected.

Liking analysis

After the analysis of WTP, the following section will discuss the results regarding participant's ratings of liking for the four different labels. Liking was rated on a 5-point scale by means of a smiley. The face of the smiley changed from very unlikable (a score of 1) to very likeable (a score of 5). It is important to mention that scores thus range between 1 and 5.

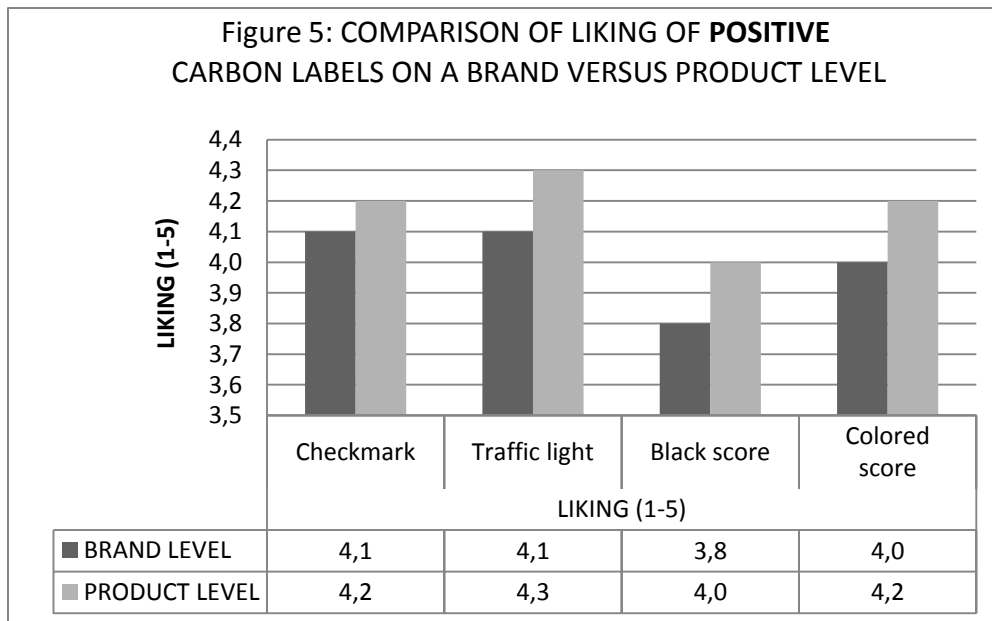
Liking Brand versus Product level, positive label

After the analysis of WTP, the following analysis will be conducted with participant's liking for fridges with a positive carbon label on a brand and on a product level as dependent variable. Condition and NEP were the predictors, and perceptions of influence and social norms were taken as covariates. This analysis reveals the following:

Within subjects

The Tests of Within-Subjects Effects shows a significant main effect ($F(1, 121) = 9.834, p = .002$) of the information on the positive carbon label on a brand level ($M = 3.97, SD = .61$) versus on a product level ($M = 4.16, SD = .65$). This effect tells us that, if we ignore the fact whether the rating was performed for four different types of labels, the ratings of liking for fridges with a carbon label on a brand level differed significantly from the ratings of liking for fridges with a carbon label on a product level. It can be seen that people rate their liking higher for a fridge with a positive carbon label on a product level as compared to a fridge with a positive carbon label on a brand level. This finding is in contrast with hypothesis 8.

Condition had no main effect ($F(3,121) = .050, p = .985$) on ratings of liking between brand versus product level. Mean differences in ratings of liking between the four different positive carbon labels on a brand versus product level are displayed in figure 5. This finding is in contrast with hypothesis 3.



Furthermore, NEP value had a significant main effect on liking for fridges with a carbon label on a brand level versus a carbon label on a product level, $F(1,121) = 13.574$, $p = .000$. This shows that the ratings for liking between respondents with a high or low NEP significantly differed from each other. Participants with a low NEP rating the liking of a fridge with a carbon label on a brand level on average the same ($M = 3.97$, $SD = .58$) as participants with a high NEP ($M = 3.97$, $SD = .64$). For a fridge with a carbon label on a product level were low NEP participants giving lower rates of liking ($M = 3.98$, $SD = .73$) as compared to participants with a high NEP ($M = 4.34$, $SD = .51$). This is in contrast with hypothesis 11.

No significant effects of the controlling covariates perceptions of influence ($p = .182$) and social norms ($p = .105$) were found.

Between subjects

Regarding the different labels, there was a significant simple main effect of condition, $F(3,121) = 2.988$, $p = .034$. This effect tells us that participants gave a different liking score between the four different labels (check mark, traffic light, black score or colored score). The mean differences of liking between the different labels can be found in table 4. Black score ($M = 3.81$, $SE = .10$) differed significantly from checkmark ($M = 4.14$, $SE = .10$). This implies that participants liked a checkmark on average .33 higher as compared to the black score label ($p = .097$). Black score ($M = 3.81$, $SE = .10$) also differed significantly from traffic light ($M = 4.17$, $SE = .09$). This implies that participants liked a traffic light with on average .36 more as compared to the black score label ($p = .053$).

Colored score ($M = 4.11$, $SE = .10$) was rated with an average liking a little less as compared to checkmark and traffic light. This shows that labels with a clear reference point (checkmark and traffic light) are preferred over a black score. However, these labels are almost similarly liked as the colored score label. This is not in line with hypothesis 3.

Table 4. Pairwise comparison labels

(I)Condition	(J)Condition	Mean Difference (I-J)	Std. Error	Sig.*	95% Confidence interval for Difference*	
					Lower bound	Upper Bound
Checkmark	Traffic light	-.022	.133	1.000	-.387	.335
	Black Score	.332	.136	.097	-.033	.679
	Colored Score	.039	.135	1.000	-.323	.401
Traffic light	Checkmark	.022	.133	1.000	-.335	.378
	Black Score	.345	.133	.053	-.002	.710
	Colored Score	.061	.133	1.000	-.295	.416
Black Score	Checkmark	-.332	.136	.097	-.697	.033
	Traffic light	-.354	.133	.053	-.710	.002
	Colored Score	-.293	.135	.195	-.656	.070
Colored Score	Checkmark	-.039	.135	1.000	-.401	.323
	Traffic light	-.061	.133	1.000	-.416	.295
	Black Score	.293	.135	.195	-.070	.656

*. Adjustment for multiple comparisons: Bonferroni

NEP had a significant simple main effect between subjects $F(1,121) = 4.227$, $p = 0.042$. This effect tells us that participants' NEP had an effect on their liking. Bonferroni corrected post hoc tests showed that ratings of liking between respondents who scored high ($M = 4.16$, $SE = .07$) versus low ($M = 3.96$, $SE = .07$) on NEP did significantly differ ($p = .042$). Respondents with high NEP score rated liking on average .20 higher for a label as compared to respondents who scored low on NEP.

No significant effects of the controlling covariates perceptions of influence ($p = .734$) and social norms ($p = .276$) were found.

Finally, a simple main effect of brand versus product was found. A significant difference could be found between the liking for a fridge with a positive carbon label on a brand level ($M = 3.96$, $SE = .05$) versus a fridge with a positive carbon label on a product level

($M = 4.15$, $SE = .06$). This implies that participants rated their liking for a fridge with a carbon label on a product level higher as compared to a fridge with a carbon label on a brand level ($Mean\ difference = .19$, $p = .001$). This is in contrast with hypothesis 8.

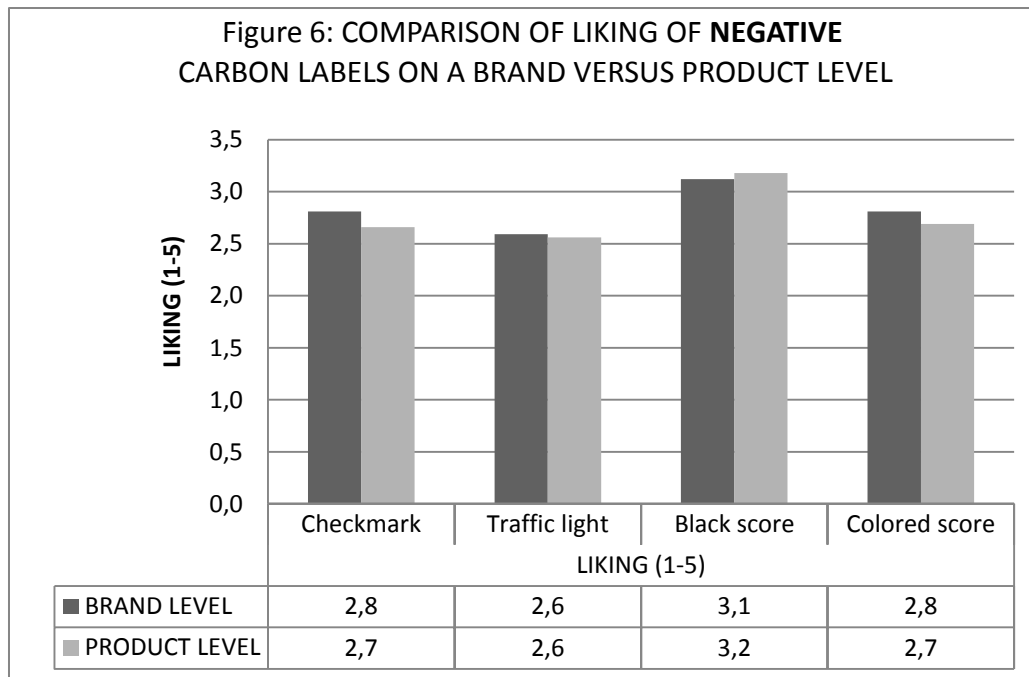
Liking Brand versus Product level, negative label

The same analysis was conducted for liking, with negative carbon labels on a brand and on a product level as dependent variable. Condition and NEP were the predictors, and perceptions of influence and social norms were taken as covariates. This analysis reveals the following:

Within subjects

No significant difference was found between the liking for fridges with a negative carbon label on a brand versus a product level, $F(1,121) = .880$, $p = .350$. Participants rated a negative carbon label on a brand level on average with 2.83 ($SD = .95$). This was 2.77 ($SD = .93$) for a negative carbon label on a product level. This is first time that a label on a brand level is rated as more positive as compared to a carbon label on a product level, but this difference is not significant. This tells us that there is no significant difference between the rated liking for fridges with a negative carbon label on a brand versus a product level, and thus that people equally “like” the fridges when the label is negative, regardless of whether the information is on a brand level or on a product level. This is in contrast with hypothesis 8.

Condition had no effect on liking for fridges with a negative carbon label on whether a brand level or on a product level ($F(3,121) = .518$, $p = .671$). This implies that fridges with a negative label were almost equally liked, regardless of whether the fridge had a label on a brand level or on a product level. Mean differences in ratings of liking between the four different negative carbon labels on a brand versus product level are displayed in figure 6.



The controlling variable social norms appeared to correlate with liking, $F(1,121) = 3.119$, $p = .080$. This influence is in a positive direction for a negative carbon label on a brand level ($B = .31$, $p = .000$) and for a negative carbon label on a product level ($B = .17$, $p = .050$).

No significant effects for the other within subjects variables were found. Perceptions of influence had no effect as a controlling covariate ($p = .799$) and neither has NEP ($p = .696$) as a predicting variable.

Between subjects

Regarding the different labels, we found a significant simple main effect of condition, $F(3,121) = 3.129$, $p = .028$. This effect tells us that participants gave a different liking score between the four different labels (check mark, traffic light, black score, and colored score). Black score ($M = 3.12$, $SE = .14$) differed significantly from traffic light ($M = 2.54$, $SE = .13$). This implies that participants liked a fridge with a black score label with a mean difference of .58 more, as compared to the fridge with a traffic light label ($p = .018$). This is in contrast with hypothesis 3.

NEP also had a significant simple main effect between subjects $F(1,121) = 2.985$, $p = .087$. This effect tells us that participants' NEP had an effect on their liking. Bonferroni corrected post hoc tests showed that ratings of liking between respondents who scored high ($M = 2.67$, $SE = .10$) versus low ($M = 2.92$, $SE = .10$) on NEP did significantly differ.

Respondents with a low NEP gave significantly higher ratings of liking (*Mean difference* = .25) as compared to respondents who scored high on NEP. This is in contrast with hypothesis 9.

The controlling variable social norms appeared to correlate with liking, $F(1,121) = 9.838$, $p = .002$. This influence is in a positive direction for a negative carbon label on a brand level ($B = .31$, $p = .000$) and for a negative carbon label on a product level ($B = .17$, $p = .050$).

The controlling variable perceptions of influence appeared to correlate with liking, $F(1,121) = 8.287$, $p = .005$. This influence is in a negative direction for a negative carbon label on a brand level ($B = -.27$, $p = .002$) and for a negative carbon label on a product level ($B = -.25$, $p = .005$).

No simple main effect of brand level versus product level was found. A non-significant difference between the liking for a fridge with a carbon label on a brand level ($M = 2.82$, $SE = .08$) versus a fridge with a carbon label on a product level ($M = 2.76$, $SE = .08$) appeared. This shows that, on average, participants did not like a fridge with a negative carbon label on a product level more as compared to a fridge with a negative carbon label on a product level ($p = .445$). This again, is in contrast with hypothesis 8.

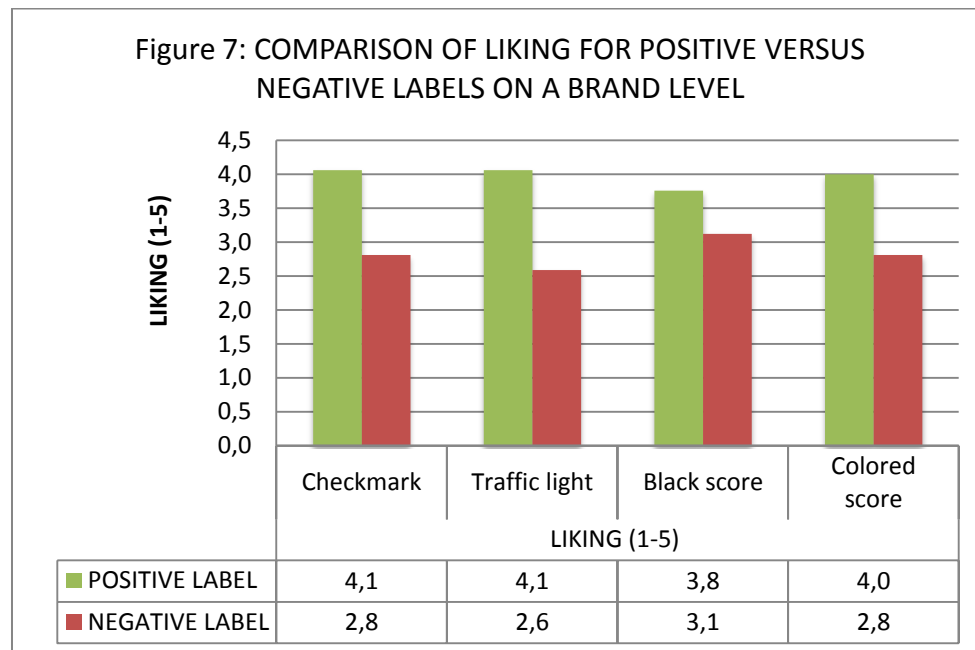
Liking Brand level, positive versus negative labels

The participants' rates of liking regarding fridges with positive and negative carbon labels on a brand level were analysed with condition and NEP as predictors, and perceptions of influence and social norms as covariates, the results reveal the following:

Within subjects

The Tests of Within-Subjects Effects shows a significant main effect of a positive ($M = 3.97$, $SD = .61$) versus a negative ($M = 2.83$, $SD = .95$) carbon label on a brand level $F(1, 121) = 4.276$, $p = .041$. This effect shows that, if we ignore the fact whether the rating was performed for different types of labels, the ratings for liking between fridges with positive and negative carbon labels on a brand level still significantly differed. It can be seen that people like a fridge with a positive carbon label on a brand level more as compared to a fridge with a negative label on a brand level. This is in line with hypothesis 2.

A significant main effect of condition was found ($F(3,121) = 3.722, p = .013$). This implies that condition has an effect on the difference between ratings of liking for a fridge with a positive versus a negative carbon label on a brand level. Mean differences in ratings of liking between the four different positive and negative carbon labels on a brand level are displayed in figure 7.



NEP value had no significant main effect with liking for positive versus negative carbon labels on a brand level, $F(1,121) = 2.035, p = .156$. This shows that participants with a low NEP ($M = 3.97, SD = .58$) versus a high NEP ($M = 3.97, SD = .64$) gave no significantly different ratings of liking for the fridge with a positive carbon label on a brand level. Neither for a fridge with a negative carbon label on a brand level differed the ratings of liking between participants with a low NEP ($M = 3.00, SD = .93$) from ratings of participants with a high NEP ($M = 2.27, SD = .96$). Here we can see a reversed effect of NEP on liking for negative labels again. Even though this effect is not significant, low NEP participants rated negative labels higher as compared to high NEP participants, when the label was on a brand level. Liking for positive labels on a brand level was ranked the same among high versus low NEP participants. This is in contrast with hypothesis 9.

The controlling variable Perceptions of influence appeared to correlate with liking, $F(1,121) = 7.665, p = .007$. This influence is in a positive direction for the positive carbon

label on a brand level ($B = .12, p = .175$), while it is in a negative direction for a negative carbon label on a brand level ($B = -.27, p = .002$).

The controlling variable Social norms appeared to correlate with liking, $F(1,121) = 3.988, p = .048$. This influence is in a positive direction for both a positive carbon label on a brand level ($B = .14, p = .11$) as well as for a negative carbon label on a brand level ($B = .31, p = .000$).

Between subjects

Condition appeared not to have a significant simple main effect on liking for a positive versus a negative carbon label on a brand level ($F(3,121) = .409, p = .746$). This implies that the ratings of liking between the four labels (checkmark, traffic light, black score, and colored score) was about the same. This is in contrast with hypothesis 3.

No significant simple main effect of NEP on the ratings of liking for a positive versus negative label on a brand level was found ($F(1,121) = 1.927, p = .168$). Bonferroni corrected post hoc tests showed that ratings of liking between respondents who scored high ($M = 3.32, SE = .07$) versus low ($M = 3.46, SE = .07$) on NEP did not significantly differed from each other ($p = .168$). Respondents with a high NEP score rated their liking on average the same as rating of people with a low NEP. This is in contrast with hypothesis 9.

Perceptions of influence ($p = .109$) as controlling variable did not have any effect.

The controlling variable Social norms appeared to correlate with liking, $F(1,121) = 15.513, p = .000$. This influence is in a positive direction for both a positive carbon label on a brand level ($B = .14, p = .106$) as well as for a negative carbon label on a brand level ($B = .31, p = .000$).

Finally, a simple main effect of positive versus negative carbon label on a brand level was found. A significant difference between the liking for a positive carbon label on a brand level ($M = 3.96, SE = .05$) versus a negative carbon label on a brand level ($M = 2.82, SE = .08$) was shown. This shows that participants liked a fridge with a positive carbon label on a brand level more (*Mean difference* = 1.14) as compared to a fridge with a negative label on a brand level ($p = .000$). This is in line with hypothesis 2.

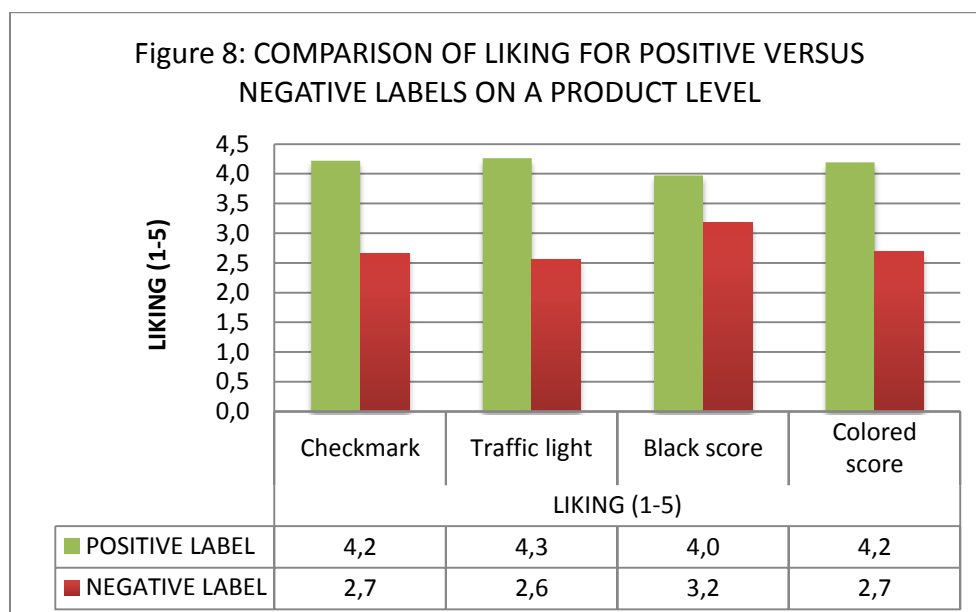
Liking Product level, positive versus negative labels

Additionally, an analysis with liking for fridges with positive and negative carbon labels on a product level was conducted. In this case, condition and NEP were still predictors, and perceptions of influence and social norms covariates. This analysis reveals the following:

Within subjects

The Tests of Within-Subjects Effects shows a significant main effect of a positive carbon label ($M = 4.16$, $SD = .65$) versus a negative carbon label ($M = 2.77$, $SD = .93$) on a product level $F(1, 121) = 9.398$, $p = .003$. This effect tells us that, if we ignore the fact whether the rating was performed for different types of labels, the ratings of liking for fridges with a positive carbon label on a product level still significantly differed from ratings of liking for fridges with a negative carbon label on a product level. This shows us that participants like a fridge with a positive carbon label on a product level more as compared to a fridge with a negative carbon label on a product level. This is in line with hypothesis 2.

A significant main effect of condition was found ($F(3,121) = 5.665$, $p = .001$). This implies that condition has an effect on the difference between ratings of liking for a fridge with a positive versus a negative carbon label on a product level. Mean differences in ratings of liking between the four different positive and negative carbon labels on a product level are displayed in figure 8.



NEP had a significant main effect on ratings of liking for positive versus negative carbon labels on a product level, $F(1,121) = 10.126$, $p = .002$. This shows that participants' ratings of liking significantly differed. Participants with a low NEP gave lower ratings of liking ($M = 3.98$, $SD = .73$) as compared to high NEP respondents ($M = 4.34$, $SD = .51$) for a fridge with a positive label on a product level. Also for a fridge with a negative carbon label differed the rating of liking between respondents with a low NEP ($M = 2.89$, $SD = .88$) versus a high NEP ($M = 2.65$, $SD = .96$) significantly. Here we see a reversed effect of NEP on liking for a negative label again. Participants with a high NEP liked a positive label more as compared to low NEP participants, and low NEP participants gave higher ratings of liking for fridges with a negative label as compared to high NEP respondents. This is not in line with hypothesis 9.

The controlling variable Perceptions of influence appeared to correlate with liking, $F(1,121) = 3.190$, $p = .077$. This influence is in a positive direction for the positive carbon label on a product level ($B = .16$, $p = .070$), while it is in a negative direction for a negative carbon label on a product level ($B = -.09$, $p = .311$).

The controlling variable Social norms appeared not to correlate with liking, $F(1,121) = 2.170$, $p = .143$.

Between subjects

Condition appeared not to have a significant simple main effect on the liking for fridges with positive or negative carbon labels on a product level, ($F(3,121) = .270$, $p = .847$). This implies that the ratings of liking between the four labels (checkmark, traffic light, black score, and colored score) were about the same. This is in contrast with hypothesis 3.

As a result of the between subjects test, NEP appeared not to have a significant simple main effect on the liking for a positive versus negative label on a product level ($F(1,121) = .880$, $p = .350$). Bonferroni corrected post hoc tests showed that ratings of liking between respondents who scored high ($M = 3.51$, $SE = .07$) versus low ($M = 3.41$, $SE = .07$) on NEP did not significantly differed from each other. Respondents with a high NEP score rated their liking on average the same as rating of people with a low NEP. This is in contrast with hypothesis 10.

The controlling variable Perceptions of influence appeared to correlate with liking, $F(1,121) = 4.401$, $p = .038$. This influence is in a positive direction for the positive carbon label on a product level ($B = .16$, $p = .070$), while it is in a negative direction for a negative carbon label on a product level ($B = -.09$, $p = .311$).

The controlling variable Social norms did not correlate with liking, $F(1,121) = 2.510$, $p = .116$.

Finally, a simple main effect of positive versus negative carbon label on a product level was found. A significant difference between the liking for a positive carbon label on a product level ($M = 4.15$, $SE = .06$) versus a negative carbon label on a product level ($M = 2.76$, $SE = .08$) was shown. This shows that participants liked a fridge with a positive carbon label on a product level more (*Mean difference* = 1.39) as compared to a fridge with a negative label on a product level ($p = .000$). This is in line with hypothesis 2.

Conclusion liking

Based on the findings in the previous sections, we can conclude the following:

H2: Positive labels (green checkmark, green traffic light, a black score above 66 and a green score above 66) will be preferred over negative labels (red checkmark, red traffic light, a black score under 33 and a red score under 33).

A significant difference between ratings of liking for positive carbon labels versus negative carbon labels was found. Positive carbon labels were liked more compared to negative carbon labels. This is in line with **hypothesis 2**, which will be **accepted**.

H3: Carbon emission labels containing a checkmark and a traffic light label will be preferred over a score label (black or colored).

Condition appeared not to have a effect on liking. That the ratings of liking between the four labels (checkmark, traffic light, black score, and colored score) was about the same, regardless of label being on a brand or product level. This is in contrast with **hypothesis 3**, which as a result will be **rejected**.

H8: Respondents will give higher ratings of liking and attractiveness for labels on a brand level compared to labels on a product level.

No significant difference was found between the liking for fridges with a carbon label on a brand versus a product level. Participants rated a carbon label on a brand level on average the same as a carbon label on a product level. This was contradicting the expectations of **hypothesis 8**, which thus will be **rejected**.

H9: NEP value will influence participants' attitude towards the products in a way that high NEP value participants will prefer environmental friendly products, where low NEP participants will make no distinction between environmental friendly and unfriendly products.

Previous section shows that that high NEP participants prefer environmental friendly products, but it cannot be stated that low NEP participants make no distinction between environmental friendly and unfriendly products. Therefore **hypothesis 9** is **rejected**

10: NEP value will influence participants' preference of labels in a way that low NEP value participants will prefer labels with a reference point (checkmark, traffic light), while there will be no difference in preference for the labels among high NEP participants.

NEP appeared not to have a significant effect on the liking scores for the different labels. No difference was found in ratings of high versus low NEP participants regarding the different types of labels. This is in contrast with **hypothesis 10**, which as a result will be **rejected**.

H11: Low NEP value participants will prefer a label on a brand level over a label on a product label, while high NEP value participants will make no distinction between a brand versus a product level, but between positivity and negativity of the label.

NEP value had no effect as a covariate on liking rates of a label on a brand level versus on a product level. Ratings of high NEP value participants were similar towards ratings of low NEP value participants. This is in contrast with **hypothesis 11**, which thus will be **rejected**.

Taken this together, it appears that hypotheses 3, 8, 9, 10, and 11 are rejected. Hypothesis 2 is accepted.

Attractiveness analysis

After the analysis of liking, the following section will discuss the results regarding participant's ratings of attractiveness for fridges with different labels. Attractiveness was rated on a 7-point Likert-scale, in which was 1 coded as "very unattractive" and 7 as "very attractive".

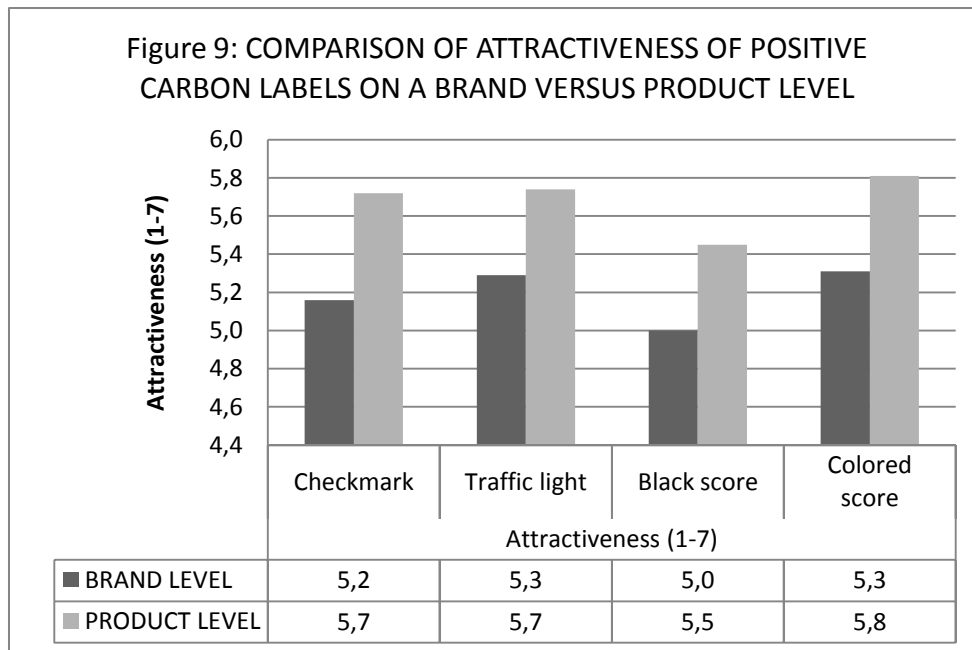
Attractiveness Brand versus Product level, positive label

The following analysis will be conducted with participant's ratings of attractiveness for fridges with a positive carbon label on a brand and on a product level as dependent variable. Condition and NEP were the predictors, and perceptions of influence and social norms were taken as covariates. This analysis reveals the following:

Within subjects

The Tests of Within-Subjects Effects shows a significant main effect ($F(1, 121) = 14.265$, $p = .000$) of the information on the positive carbon label on a brand level ($M = 5.19$, $SD = 1.11$) versus on a product level ($M = 5.68$, $SD = 1.05$). This effect tells us that, if we ignore the fact whether the rating was performed for four different types of labels, the ratings of attractiveness for fridges with a carbon label on a brand level differed significantly from the ratings of attractiveness for fridges with a carbon label on a product level. It can be seen that people rate their attractiveness higher for a fridge with a positive carbon label on a product level as compared to a fridge with a positive carbon label on a brand level. This is in contrast with hypothesis 8.

Condition had no main effect ($p = .715$) on the attractiveness ratings between brand versus product. Mean differences in ratings of attractiveness between the four different positive carbon labels on a brand versus product level are displayed in figure 9.



NEP had a significant main effect on attractiveness for fridges with a carbon label on a brand level versus a carbon label on a product level, $F(1,121) = 9.050$, $p = .003$. This shows that the ratings for attractiveness between respondents with a high or low NEP significantly differed from each other. Participants with a low NEP rated the attractiveness of a fridge with a carbon label on a brand level as less attractive ($M = 5.14$, $SD = 1.08$) as compared to participants with a high NEP ($M = 5.25$, $SD = 1.15$). Participants with a low NEP rated the attractiveness of a fridge with a carbon label on a product level also less ($M = 5.42$, $SD = 1.07$) as compared to participants with a high NEP ($M = 5.94$, $SD = .97$). This is in contrast with hypothesis 11.

The controlling variable perceptions of influence appeared to correlate with attractiveness, $F(1,121) = 4.270$, $p = .041$. This influence is in a positive direction for the positive carbon label on a brand level ($B = .10$, $p = .293$) and in a negative direction for the positive carbon label on a product level ($B = -.01$, $p = .935$).

No significant effect of the controlling covariate social norms ($p = .292$) was found.

Between subjects

Regarding the different labels, there was no significant simple main effect of condition on ratings of attractiveness ($p = .285$). This effect tells us that participants gave no different ratings of attractiveness towards the fridges between conditions, which are the four

different labels (check mark, traffic light, black score or colored score). This is in contrast with hypothesis 3.

NEP had a significant simple main effect between subjects ($F(1,121) = 4.002, p = .048$). This effect tells us that participants' NEP had an effect on their attractiveness towards the fridges. Bonferroni corrected post hoc tests showed that ratings of attractiveness between respondents who scored high ($M = 5.60, SE = .11$) versus low ($M = 5.26, SE = .12$) on NEP did significantly differ. Respondents with high a NEP score rated the attractiveness of products with a positive label higher (*Mean difference* = .34) as compared to respondents who scored low on NEP. This is again in contrast with hypothesis 11.

The controlling covariate social norms appeared to correlate with attractiveness, $F(1,121) = 3.070, p = .082$. This influence is in a positive direction for the positive carbon label on a brand level ($B = .10, p = .283$) and in a negative direction for the positive carbon label on a product level ($B = -.01, p = .935$).

Perceptions of influence had no significant effect as a controlling covariate ($p = .841$).

Finally, a simple main effect of brand level versus product level was found. A significant difference could be found between the rates of attractiveness for a fridge with a positive carbon label on a brand level ($M = 5.19, SE = .10$) versus a fridge with a positive carbon label on a product level ($M = 5.67, SE = .09$). This shows that participants rated the attractiveness of a fridge with a carbon label on a product level higher as compared to a fridge with a carbon label on a brand level (*Mean difference* = .48, $p = .000$). This is contradicting to hypothesis 8.

Attractiveness Brand versus Product level, negative label

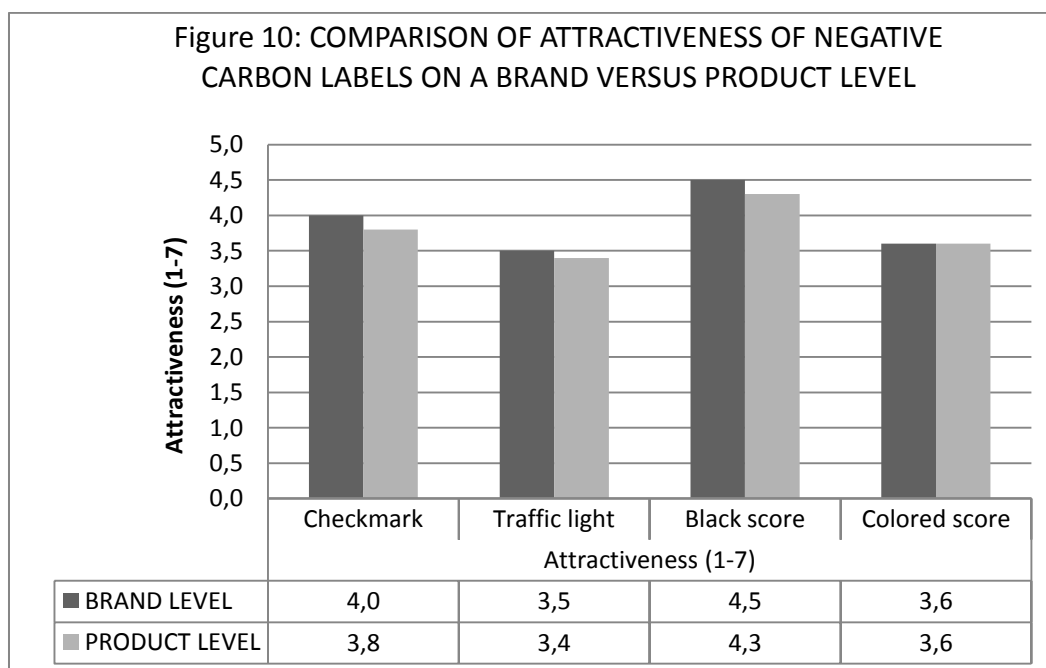
The same analysis was conducted for attractiveness, with negative carbon labels on a brand level and on a product level as dependent variable. Condition and NEP were the predictors, and perceptions of influence and social norms were taken as covariates. This analysis reveals the following:

Within subjects

No significant difference was found between ratings of attractiveness for fridges with a negative carbon label on a brand level versus on a product level ($F(1,121) = .116$, $p = .734$). Participants rated the attractiveness of a fridge with a negative carbon label on a brand level on average with 3.87 ($SD = 1.36$) as compared to 3.76 ($SD = 1.42$) for a fridge with a negative carbon label on a product level. This tells us that there is no significant difference between the rated attractiveness for fridges with a negative carbon label on a brand level versus a product level, and thus that people are equally “attracted” to the fridges when the label is negative, regardless of whether the information is on a brand level or on a product level. This is in contrast with hypothesis 8.

Furthermore, no significant effects within subjects were found. Perceptions of influence ($p = .548$) and social norms ($p = .637$) had no effect as a controlling covariates. Neither NEP had a significant main effect on ratings of attractiveness ($p = .637$).

Condition had no main effect ($p = .851$) on the attractiveness ratings between brand versus product. Mean differences in ratings of attractiveness between the four different negative carbon labels on a brand versus product level are displayed in figure 10.



Between subjects

Regarding the different labels, there was a significant simple main effect of condition, $F(3,121) = 4.122$, $p = .008$. This effect tells us that participants gave a different score of

attractiveness between the four different labels (check mark, traffic light, black score, and colored score). Black score ($M = 4.29$, $SE = .20$) differed significantly from traffic light ($M = 3.39$, $SE = .19$). This implies that participants rated the attractiveness of a fridge with a black score label on average with .90 higher as compared to the fridge with a traffic light label ($p = .007$). Ratings of attractiveness for fridges with a black score label also differed significantly from ratings of attractiveness for a fridge with a colored score label ($M = 3.62$, $SE = .19$). This implies that participants rated attractiveness of fridges with a black score on average with .67 higher as compared to the fridges with a colored score label ($p = .098$). Checkmark label ($M = 3.94$, $SE = .20$) did not significantly differ from the other labels. Yet, it is still visible that the mean score of checkmark label ($M = 3.94$, $SE = .20$) is lower than the mean of the black score label ($M = 4.29$, $SE = .20$). This shows that colored score and traffic light labels are ranked as significantly less attractive when the label is negative. This is opposed to the findings regarding positive labels. However, these findings are in contrast with hypothesis 3, which states that the black score will be seen as the least attractive.

NEP had no significant simple main effect between subjects $F(1,121) = .792$, $p = .375$. This effect tells us that participants' NEP had no effect on their ranking of attractiveness, which is in contrast with hypothesis 11.

The controlling variable perceptions of influence appeared to correlate with attractiveness, $F(1,121) = 14.699$, $p = .000$. This influence is in a negative direction for a negative carbon label on a brand level ($B = -.31$, $p = .000$) and for a negative carbon label on a product level ($B = -.29$, $p = .001$).

The controlling variable social norms appeared to correlate with attractiveness, $F(1,121) = 19.850$, $p = .000$. This influence is in a positive direction for a negative carbon label on a brand level ($B = .35$, $p = .000$) as well as for a negative carbon label on a product level ($B = .30$, $p = .000$).

No simple main effect of brand level versus product level was found. A non-significant difference could be found between the ratings of attractiveness for a fridge with a carbon label on a brand level ($M = 3.86$, $SE = .11$) versus a fridge with a carbon label on a product level ($M = 3.76$, $SE = .12$). This shows that, on average, participants were not more or less attracted towards a fridge with a negative carbon label on a product level as

compared to a fridge with a negative carbon label on a product level ($p = .374$). This is in contrast with hypothesis 3.

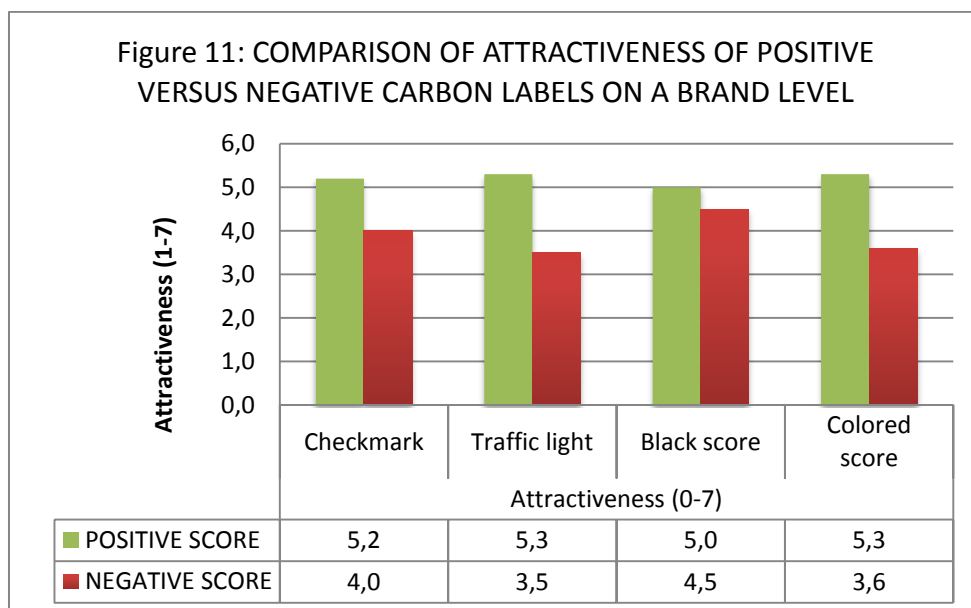
Attractiveness Brand level, positive versus negative labels

Participant's attractiveness regarding fridges with positive and negative carbon labels on a brand level was analysed with condition and NEP as predictors, and perceptions of influence and social norms as covariates, the results reveal the following:

Within subjects

The Tests of Within-Subjects Effects showed no significant main effect of a positive ($M = 5.19$, $SD = 1.11$) versus a negative ($M = 3.87$, $SD = 1.36$) carbon label on a brand level, $F(1, 121) = .830$, $p = .364$. This effect shows that, if we ignore the fact whether the rating was performed for different types of labels, the ratings of attractiveness between fridges with positive and negative carbon labels on a brand level did not significantly differed. This is in contrast with hypothesis 2.

A significant main effect of condition was found ($F(3,121) = 5.010$, $p = .003$). This implies that type of label had an effect on the difference between ratings of attractiveness for a fridge with a positive versus a negative carbon label on a brand level. Mean differences in ratings of attractiveness between the four different positive versus negative carbon labels on a brand level are displayed in figure 11.



NEP value had no significant main effect with attractiveness for positive versus negative carbon labels on a brand level, $F(1,121) = .590$, $p = .444$. This shows that participants with a low NEP ($M = 5.14$, $SD = 1.08$) versus a high NEP ($M = 5.25$, $SD = 1.15$) gave no significantly different ratings of attractiveness for the fridge with a positive carbon label on a brand level. Neither for a fridge with a negative carbon label on a brand level differed the ratings of liking between participants with a low NEP ($M = 4.00$, $SD = 1.32$) versus ratings of attractiveness of participants with a high NEP ($M = 3.74$, $SD = 1.40$). In line with findings regarding rates of liking for labels on a brand level, here we can see a reversed effect of NEP on attractiveness for negative labels again. Even though this effect is not significant, low NEP participants rated the fridges with negative labels as more attractive as compared to high NEP participants, when the label was on a brand level. Attractiveness for fridges with positive labels on a brand level was ranked the same among high versus low NEP participants. This is in contrast with hypothesis 9.

A significant interaction effect of condition * NEP on the difference between attractiveness of a positive carbon label versus a negative carbon label on a brand level ($F(3,121) = 2.409$, $p = .070$). This shows us that there is a significant difference between ratings of attractiveness for a positive and negative carbon label per condition, taken into account the NEP of participants.

The controlling variable perceptions of influence appeared to correlate with attractiveness, $F(1,121) = 13.014$, $p = .000$. This influence is in a positive direction for the positive carbon label on a brand level ($B = .10$, $p = .283$), while it is in a negative direction for a negative carbon label on a brand level ($B = -.31$, $p = .000$).

The controlling variable social norms appeared to correlate with attractiveness, $F(1,121) = 4.051$, $p = .046$. This influence is in a positive direction for the positive carbon label on a brand level ($B = .16$, $p = .068$) as well as for a negative carbon label on a brand level ($B = .35$, $p = .000$).

Between subjects

Condition appeared not to have a significant simple main effect on the ratings of attractiveness for a positive versus negative carbon label on a brand level ($F(3,121) = .743$, $p = .529$). So, participants did not give a different rating for attractiveness between the different conditions. This is in contrast with hypothesis 3.

Neither NEP appeared to have a significant simple main effect on the ratings of attractiveness for a positive versus negative carbon label on a brand level ($F(1,121) = .016, p = .898$). Bonferroni corrected post hoc tests showed that ratings of attractiveness between respondents who scored high ($M = 4.51, SE = .11$) versus low ($M = 4.53, SE = .11$) on NEP did not significantly differ ($p = .898$). Respondents with a high NEP score rated their attractiveness for the fridges on average the same as rating of people with a low NEP. This finding is not in line with hypothesis 9.

The controlling variable Perceptions of influence appeared to correlate with attractiveness, $F(1,121) = 4.562, p = .035$. This influence is in a positive direction for the positive carbon label on a brand level ($B = .10, p = .283$), while it is in a negative direction for a negative carbon label on a brand level ($B = -.31, p = .000$).

The controlling variable Social norms appeared to correlate with attractiveness, $F(1,121) = 17.329, p = .000$. This influence is in a positive direction for the positive carbon label on a brand level ($B = .16, p = .068$) as well as for a negative carbon label on a brand level ($B = .35, p = .000$).

Finally, a simple main effect of positive versus negative carbon label on a brand level was found. A significant difference between the attractiveness of a fridge with a positive carbon label on a brand level ($M = 5.19, SE = .10$) versus a fridge with a negative carbon label on a brand level ($M = 3.86, SE = .11$) was shown. This shows that participants were more attracted to a fridge with a positive carbon label on a brand level ($Mean\ difference = 1.33$) as compared to a fridge with a negative label on a brand level ($p = .000$). This is in line with hypothesis 2.

Attractiveness product level, positive versus negative label

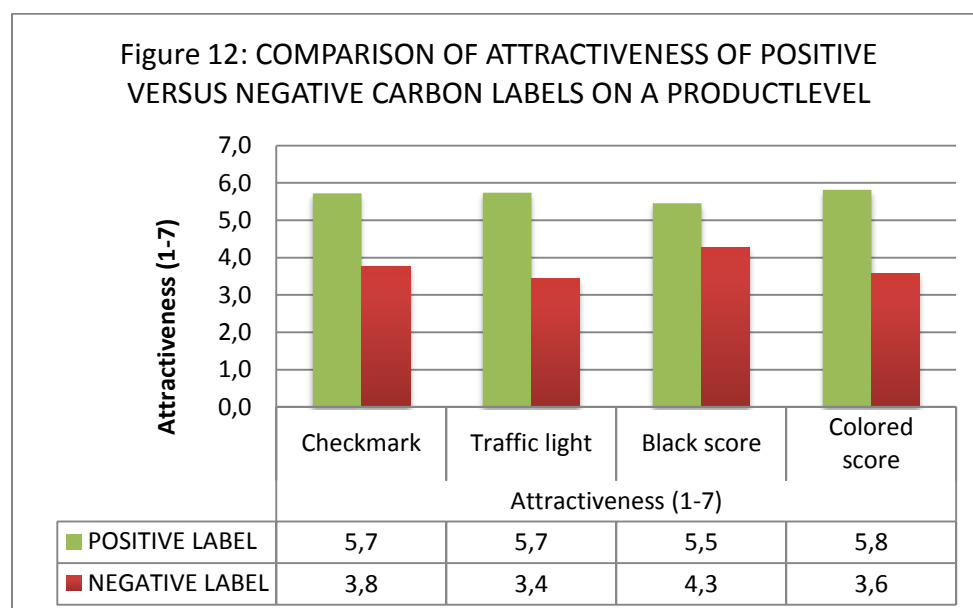
Additionally, an analysis with attractiveness for fridges with positive and negative carbon labels on a product level was conducted. In this case, condition and NEP were still predictors, and perceptions of influence and social norms covariates. This analysis reveals the following:

Within subjects

The Tests of Within-Subjects Effects shows a significant main effect of a positive carbon label ($M = 5.68, SD = 1.05$) versus a negative carbon label ($M = 3.76, SD = 1.42$) on a

product level $F(1, 121) = 10.969, p = .001$. This effect tells us that, if we ignore the fact whether the rating was performed for different types of labels, the ratings of attractiveness for fridges with a positive carbon label on a product level still significantly differed from ratings of liking for fridges with a negative carbon label on a product level. This shows us that participants are more attracted to a fridge with a positive carbon label on a product level more as compared to a fridge with a negative carbon label on a product level. This is in line with hypothesis 2.

A significant main effect of condition was found ($F(3,121) = 3.889, p = .011$). This implies that type of label has an effect on the difference between ratings of attractiveness for a fridge with a positive versus a negative carbon label on a product level. Mean differences in ratings of attractiveness between the four different positive versus negative carbon labels on product level are displayed in figure 12.



NEP had a significant main effect on ratings of attractiveness for positive versus negative carbon labels on a product level, $F(1,121) = 8.003, p = .005$. This shows that participants' ratings of attractiveness significantly differed. Participants with a low NEP gave lower ratings of attractiveness ($M = 5.42, SD = 1.07$) as opposed to participants with a high NEP ($M = 5.94, SD = .97$) for a fridge with a positive label on a product level. Also for a fridge with a negative carbon label differed the ratings of attractiveness between low NEP participants ($M = 3.92, SD = 1.36$) versus high NEP participants ($M = 3.58, SD =$

1.47). Here we see a reversed effect of NEP on attractiveness for a negative carbon label again, in which low NEP participants rate the attractiveness of negative labels higher as compared to high NEP participants. As a result, this is not supporting hypothesis 9.

The controlling variable Perceptions of influence did not appear to correlate with attractiveness, $F(1,121) = 2.635$, $p = .107$.

The controlling variable Social norms correlates with attractiveness, $F(1,121) = 4.431$, $p = .037$. This influence is in a positive direction for both the positive carbon label on a product level ($B = .07$, $p = .427$) and the negative carbon label on a product level ($B = .30$, $p = .000$).

Between subjects

Condition appeared not to have a significant simple main effect on the attractiveness for a positive versus negative label on a product level ($F(3,121) = .517$, $p = .671$). This implies that the ratings of attractiveness between the four labels (checkmark, traffic light, black score, and colored score) were about the same. This shows no difference in preference, and thus hypothesis 3 is not supported.

As a result of the between subjects test, NEP appeared not to have a significant simple main effect on the attractiveness for a fridge with a positive versus negative carbon label on a product level ($F(1,121) = 1.517$, $p = .221$). Bonferroni corrected post hoc tests showed that ratings of attractiveness between respondents who scored high ($M = 4.81$, $SE = .11$) versus low ($M = 4.62$, $SE = .11$) on NEP did not significantly differ ($p = .221$). Respondents with a high NEP score rated the attractiveness of the fridges on average the same as rating of people with a low NEP. This is in contrast with hypothesis 11.

The controlling variable perceptions of influence appeared to correlate with attractiveness, $F(1,121) = 8.800$, $p = .004$. This influence is in a negative direction for both the positive carbon label on a product level ($B = -.01$, $p = .935$) and the negative carbon label on a product level ($B = -.29$, $p = .001$).

The controlling variable Social norms appeared to correlate with attractiveness, $F(1,121) = 11.567$, $p = .001$. This influence is in a positive direction for both the positive carbon

label on a product level ($B = .07$, $p = .427$) and the negative carbon label on a product level ($B = .30$, $p = .000$).

Finally, a simple main effect of positive versus negative carbon label on a product level was found. A significant difference could be found between the attractiveness for a fridge with a positive carbon label on a product level ($M = 5.67$, $SE = .09$) versus a fridge with a negative carbon label on a product level ($M = 3.76$, $SE = .12$). This shows that participants rated a fridge with a positive carbon label on a product level as more attractive (*Mean difference* = 1.91) as compared to a fridge with a negative label on a product level ($p = .000$). This is in line with hypothesis 2.

Conclusion attractiveness

Based on the findings in the previous sections, we can conclude the following:

H2: Positive labels (green checkmark, green traffic light, a black score above 66 and a green score above 66) will be preferred over negative labels (red checkmark, red traffic light, a black score under 33 and a red score under 33).

Positive carbon labels are indicated as more attractive compared to negative carbon labels. Also when the label was a negative label with carbon information on a brand level, it was rated as less attractive compared to a positive label with information on a brand level. These findings support our assumption made in **hypothesis 2**, which as a result will be **accepted**.

H3: Carbon emission labels containing a checkmark and a traffic light label will be preferred over a score label (black or colored).

Condition (type of label) appeared not to have a effect on the ratings of attractiveness for a positive versus negative carbon label on a brand level, and neither for brand versus product level. This implies that participants did not give different ratings regarding attractiveness for the different types of labels. This is in contrast with **hypothesis 3**, which as a result will be **rejected**.

H8: Respondents will give higher ratings of liking and attractiveness for labels on a brand level compared to labels on a product level.

No significant difference was found between ratings of attractiveness for fridges with a positive or negative carbon label on a brand level versus on a product level. This shows that brand level information is not preferred over product level information, which is in contrast with hypothesis 8. As a result, **hypothesis 8** will be **rejected**.

H9: NEP value will influence participants' attitude towards the products in a way that high NEP value participants will prefer environmental friendly products, where low NEP participants will make no distinction between environmental friendly and unfriendly products.

Based on previous findings, we can state that high NEP participants prefer environmental friendly products, but it cannot be stated that low NEP participants make no distinction between environmental friendly and unfriendly products. Therefore, **hypothesis 9** will be **rejected**.

H11: Low NEP value participants will prefer a label on a brand level over a label on a product label, while high NEP value participants will make no distinction between a brand versus a product level, but between positivity and negativity of the label.

Previous findings show that respondents with a high NEP score rated the attractiveness of the fridges on average the same as rating of people with a low NEP, regardless of the label information being on a brand level or on a product level. Furthermore, high NEP participants appeared to give higher ratings towards brand level information labels. This is in contrast with **hypothesis 11**, which will be **rejected**.

Taken this all together, we can state that hypothesis 2 is accepted. Hypotheses 3, 8, 9, and 11 are rejected.

Recall analysis

Recall exact brand versus product level, positive label

In this test, recall of carbon labels with a positive rating on a brand level and on a product level were the dependent variables. Condition and NEP were the predictors, and perceptions of influence and social norms were taken as covariates. The results shows us the following:

Within subjects

There is a significant difference between the recall of positive carbon labels on a brand level versus a product level ($F(1,121) = 9.05, p = .003$). Brand level ($M = .42, SD = .50$) versus a product level ($M = .50, SD = .50$) significantly differed from each other. This shows that, if we ignore the fact whether the rating was performed for different types of labels, the recall between labels on a product and brand level still significantly differed. It can be seen that people recall a positive carbon label on a brand level better as compared to a positive label on a product level.

Condition had no significant main effect on the recall of carbon labels on a brand versus a product level, $F(1,121) = 1.563, p = .202$. This shows that the recall of labels between respondents not differed per condition. This is in contrast with hypothesis 4.

NEP had no significant main effect on the recall of the positive carbon labels on a brand level versus on a product level, $F(1,121) = .129, p = .720$. This shows that the recall between respondents with a high or low NEP did not significantly differed from each other. Low NEP participants recalled the score of a label on a brand level ($M = .38, SD = .50$) better as compared to participants with a high NEP ($M = .46, SD = .50$). This effect though is not significant. Neither for a positive label on a product level did recall of low NEP respondents ($M = .50, SD = .50$) differed from high NEP respondents ($M = .49, SD = .50$).

No significant effect of the controlling covariate perceptions of influence ($p = .345$) was found.

Social norms appeared to correlate with recall, $F(1,121) = 6.212, p = .014$. This influence is in a positive direction for the positive carbon label on a brand level ($B = .17, p = .061$).

However, for on a positive carbon label on a product level, the direction is negative ($B = -.10, p = .241$).

Between subjects

Regarding the four types of labels, there was a significant simple main effect of condition for recall, $F(3,121) = 48,080, p = .000$. This shows that there was a difference in recall between the four different labels (checkmark, traffic light, black score and colored score). Bonferroni corrected post hoc test shows that recall of a checkmark label ($M = .21, SE = .05$) differed significantly from a black score label ($M = .69, SE = .05$). This implies that participants recalled a checkmark label on average better (*Mean difference* = .48) as compared to the black score label ($p = .000$). Checkmark label ($M = .21, SE = .05$) also differed significantly from colored score ($M = .82, SE = .05$). This implies that participants recalled a checkmark label better (*Mean difference* = .61) as compared to the colored score label ($p = .000$). Traffic light ($M = .14, SE = .05$) differed significantly from black score ($M = .69, SE = .05$). This implies that participants recalled a traffic light better (*Mean difference* = .55) as compared to the black score label ($p = .000$). Traffic light ($M = .14, SE = .05$) also differed significantly from colored score ($M = .82, SE = .05$). This implies that participants recalled a traffic light better (*Mean difference* = .68) as compared to the colored score label ($p = .000$). An overview of these results can be found in table 5. This shows that the checkmark and traffic light labels are better as compared to the black- and colored scored. This is in line with hypothesis 4. Table 5 shows no significant differences between the recall of the black and the colored score. This is in contrast with hypothesis 5.

Table 5. Differences in exact recall between conditions

(I) Condition	(J) Condition	Mean Difference (I-J)	Std. Error	Sig.**	95% confidence Interval for Difference**	
					Lower Bound	Upper Bound
Checkmark	Traffic light	,070	,069	1,000	-,115	,254
	Black score	-,474*	,070	,000	-,663	-,285
	Colored score	-,609*	,070	,000	-,769	-,421
Traffic light	Checkmark	-,070	,069	,1000	-,254	,115
	Black score	-,543*	,069	,000	-,727	-,359
	Colored score	-,678*	,068	,000	-,862	-,495
Black score	Checkmark	,474*	,070	,000	,285	,663
	Traffic light	,543*	,069	,000	,359	,727
	Colored score	-,135*	,070	,339	-,322	,053
Colored score	Checkmark	,609*	,070	,000	,421	,769
	Traffic light	,678*	,068	,000	,495	,862
	Colored score	,135	,070	,339	-,053	,322

*. The mean difference is significant at the ,05 level.

**. Adjustment for multiple correlations: Bonferroni.

NEP did not had a significant simple main effect between subjects $F(1,121) = .234$, $p = 0.630$. This effect tells us that participants' NEP had no effect on their recall of the labels. Bonferroni corrected post hoc tests showed that the label recall between respondents who scored high ($M = .48$, $SE = .04$) versus low ($M = .45$, $SE = .04$) on NEP did not significantly differ.

There was no significant effect of the controlling covariates social norms ($p = .278$).

The controlling variable Perceptions of influence appeared to correlate with recall, $F(1,121) = 3.630$, $p = .059$. This influence is in a negative direction for the positive carbon label on a brand level ($B = -.09$, $p = .320$), and also in a negative direction for the positive carbon label on a product level ($B = -.16$ $p = .072$).

Finally, no simple main effect of recall for brand level versus product level was found. There was no significant difference between recall of a positive carbon label on a brand

level ($M = .43$, $SE = .03$) versus a positive carbon label on a product level ($M = .51$, $SE = .04$). So, on average, participants did not recall positive carbon labels on a brand level better as compared to a positive carbon label on a product level and vice versa ($p = .108$).

Recall approximately, brand versus product level, positive label

As opposed to the previous analysis, this test was conducted after recoding participants' answers on the labels with number scores (black and colored). The results of the test between the positive carbon labels for brand versus product level show us the following:

Within subjects

There is a significant difference between the approximate recall of positive carbon labels on a brand versus a product level ($F(1,121) = 6.739$, $p = .011$). The category of a positive carbon label on a brand level was significantly better remembered ($M = .08$, $SD = .28$) as compared to the category of a positive carbon label on a product level ($M = .21$, $SD = .41$). This effect tells us that, if we ignore the fact whether the recall test was performed for different types of labels, the recall of the correct category between positive carbon labels on a product versus brand level still significantly differed. It can be seen that people recall the category of a positive label on a brand level better as compared to a positive carbon label on a product level.

Condition had no significant main effect on the recall of category of positive carbon labels on a brand versus a product level, $F(3,121) = .653$, $p = .583$. This shows that the recall of category between respondents not differed per condition.

NEP had no significant main effect on the recall of category of the positive carbon labels on a brand versus a product level, $F(1,121) = 2.148$, $p = .145$. This shows that the recall of category between respondents with a high or low NEP did not significantly differed from each other. Participants with a low NEP recalled the category of a positive carbon label on a brand level ($M = .12$, $SD = .33$) less as compared to participants with a high NEP ($M = .05$, $SD = .21$). This effect though is not significant. Neither for positive carbon labels on a product level did low NEP ($M = .20$, $SD = .40$) differed significantly with the recall of category of high NEP respondents ($M = .23$, $SD = .43$).

No significant effect of the controlling covariate perceptions of influence ($p = .453$) was found.

Social norms appeared to correlate with recall of category, $F(1,121) = 2.893$, $p = .092$. This influence is in a positive direction for the positive carbon labels on a brand level ($B = .20$, $p = .023$). However, for a positive carbon label on a product level, the direction is negative ($B = -.05$, $p = .601$).

Between subjects

Regarding the different labels, we found a significant simple main effect of condition for recall of category, $F(3,121) = 2.269$, $p = .084$. This effect tells us that there was a difference in recall of category between the different labels. However, after correcting for Bonferroni, no significant differences were found in the post hoc test. If the test was conducted again without the Bonferroni correction, significant differences between the checkmark label ($M = .21$, $SE = .04$) and the black score label ($M = .06$, $SE = .04$) were found. Interestingly, the category of the black score label was better remembered as compared to the checkmark label (*Mean difference* = .15). Furthermore, the black score label ($M = .06$, $SE = .04$) differed significantly in recall of category from the colored score label ($M = .19$, $SE = .04$). The category of the black score label was better remembered as compared to the category of the colored score label (*Mean difference* = .13).

NEP did not had a significant simple main effect between subjects $F(1,121) = .002$, $p = 0.961$. This effect tells us that participants' NEP had no effect on their recall of category of the labels. Bonferroni corrected post hoc tests showed that the label recall of category between respondents who scored high ($M = .15$, $SE = .03$) versus low ($M = .15$, $SE = .03$) on NEP did not significantly differ.

No significant effects of the controlling covariates social norms ($p = .319$) and perceptions of influence ($p = .448$) were found.

Finally, a significant simple main effect of recall of category for brand level versus product level was found. A significant difference could be found between recall of the category of a positive label on a brand level ($M = .08$, $SE = .02$) versus a positive label on a product level ($M = .22$, $SE = .04$). This shows that, on average, participants recalled the category of a positive label on a brand level better as compared to a positive label on a product level (*Mean difference* = .14, $p = .002$).

Recall exact brand versus product level, negative label

The following test was performed in order to determine respondent's exact recall of negative carbon labels. In this test, recall of negative carbon labels on a brand level and on a product level were the dependent variables. Condition and NEP were the predictors, and perceptions of influence and social norms were taken as covariates. The results shows us the following:

Within subjects

There was no significant difference between the recall of negative carbon labels on a brand versus a product level, ($F(1,121) = .839, p = .361$). This shows that there is no significant difference between the recall of negative labels on whether a brand or a product level.

Condition had no significant main effect on the recall of negative labels on a brand versus product level, $F(3,121) = .556, p = .645$. This shows that the recall of category between respondents not differed per condition.

Neither NEP had a significant main effect on the recall of the negative labels for brand versus product level, $F(1,121) = 1.281, p = .260$. This shows that the recall between respondents with a high or low NEP did not significantly differed from each other.

Perceptions of influence appeared to correlate with recall of negative labels, $F(1,121) = 3.433, p = .066$. This was in a small positive direction for the negative carbon label on a brand level ($B = .00, p = .962$), and in a negative direction for a negative carbon label on a product level ($B = -.12, p = .182$).

Social norms had no significant effect on recall as a covariate ($p = .512$).

Between subject

Regarding the different labels, a significant simple main effect of condition was found, $F(3,121) = 139.64, p = .000$. This effect tells us that there was a difference in recall between the four different labels. Post hoc test shows that recall of the checkmark label ($M = .18, SE = .04$) differed significantly from recall of the black score label ($M = .88, SE = .04$) and from a colored score label ($M = .91, SE = .38$). The participants recalled the negative checkmark label better as compared to the negative black score label (*Mean difference* = .70, $p = .000$) and as compared to the colored score label (*Mean difference*

$=.73, p = .000$). There was also a significant difference between the recall of the traffic light label ($M = .07, SE = .37$) as compared to the black score label ($M = .88, SE = .38$) and the colored score label ($M = .91, SE = .38$). This shows that the traffic light label was better remembered (*Mean difference* $= .81, p = .000$) as compared to the black score, and as compared to the colored score (*Mean difference* $= .84, p = .000$). This is again in line with hypothesis 4.

NEP did not have a significant simple main effect between subjects $F(1,121) = .002, p = .969$. This effect tells us that participants' NEP had no effect on their recall of the labels. Bonferroni corrected post hoc tests showed that the label recall between respondents who scored high ($M = .51, SE = .03$) versus low ($M = .51, SE = .03$) on NEP did not significantly differ.

No significant effects of the controlling covariates perceptions of influence ($p = .359$) and social norms ($p = .386$) were found on the recall of a negative label on whether a brand or a product level.

Finally, no simple main effect on recall of negative carbon labels on a brand level versus a product level was found. There is no significant difference between recall of a negative carbon label on a brand level ($M = .52, SE = .03$) versus a negative carbon label on a product level ($M = .50, SE = .03$). This shows that, on average, participants did not recall a negative carbon label better whether it contained information on a brand level versus a product level ($p = .666$).

Recall approximately brand versus product level, negative label

In this test, approximately recall of negative carbon labels on a brand level and on a product level were the dependent variables. Condition and NEP were the predictors, and perceptions of influence and social norms were taken as covariates. The results of the test between the negative carbon labels for brand versus product level show us the following:

Within subjects

There is no significant difference between the approximate recall of negative carbon labels on a brand versus a product level ($F(1,121) = .175, p = .676$). This effect tells us

that, if we ignore the fact whether the recall test was performed for different types of labels, the recall of the correct category between negative carbon labels on a product versus brand level did not significantly differed from each other. A negative label on a brand level was not significantly better remembered ($M = .37$, $SD = .49$) as compared to a negative label on a product level ($M = .35$, $SD = .48$).

Condition had no significant main effect on the recall of category of negative carbon labels on a brand versus a product level $F(3,121) = .401$, $p = .753$. This shows that the recall of category between respondents not differed per condition. This is in contrast with both hypotheses 4 and 5.

NEP had no significant main effect on the recall of category of the negative carbon labels on a brand versus product level, $F(1,121) = .288$, $p = .592$. This shows that the recall of category between respondents with a high or low NEP did not significantly differed from each other. Participants with a low NEP ($M = .35$, $SD = .48$) recalled the category of a negative carbon label on a brand level a little better as compared to participants with a high NEP ($M = .40$, $SD = .49$). This effect though is not significant. Neither for a negative carbon labels on a product level did low NEP ($M = .30$, $SD = .46$) significantly differed with high NEP respondents ($M = .40$, $SD = .49$).

No significant effects of the controlling covariates perceptions of influence ($p = .793$) and social norms ($p = .496$) were found.

Between subjects

Regarding the different labels, we found a significant simple main effect of condition for recall of category, $F(3,121) = 136.553$, $p = .000$. This effect tells us that there was a difference in recall of category between the different labels. These differences were between checkmark ($M = .17$, $SE = .04$) and black score ($M = .99$, $SE = .04$). On average, participants remembered a checkmark score label better as compared to a black score label (*Mean difference* = .82, $p = .000$). Furthermore, recall of the traffic light label ($M = .07$, $SE = .04$) differed significantly from the black score label ($M = .99$, $SE = .04$) and the colored score label ($M = .21$, $SE = .04$). This differences were on average .92 ($p = .000$) and .14 ($p = .057$) respectively. Finally, black score label ($M = .99$, $SE = .04$) and the colored score label ($M = .21$, $SE = .04$) differed significantly from each other. The colored score label was remembered better than the black score label (*Mean difference* = .78, $p = .000$). This shows that the black score label was often recalled

worse as compared to the other labels. These findings are partly supporting hypotheses 4 and 5.

NEP did not have a significant simple main effect between subjects $F(1,121) = .366, p = .546$. This effect tells us that participants' NEP had no effect on their recall of category of the labels. Bonferroni corrected post hoc tests showed that the label recall between respondents who scored high ($M = .35, SE = .03$) versus low ($M = .37, SE = .03$) on NEP did not significantly differ.

Social norms appeared to correlate with the recall of category of negative carbon labels, $F(1,121) = 2.898, p = .000$. This influence is in a positive direction for the negative brand level label ($B = .07, p = .414$) as well as for the negative label on a product level ($B = .13, p = .150$).

No significant effect of the controlling covariate perceptions of influence ($p = .796$) was found.

Finally, no significant simple main effect of recall of category for brand level versus product level was found. No significant difference was found between recall of category of a negative carbon label on a brand level ($M = .37, SE = .03$) versus a negative carbon label on a product level ($M = .35, SE = .03$). So, on average, participants did remember a negative carbon label on a product level a bit better as compared to negative carbon labels on a brand level. This is opposed to the results found for positive carbon labels, where brand level was reminded better. However, this difference is only small (*Mean difference* = .02) and thus is not significant ($p = .600$).

Recall exact positive versus negative labels, Brand level

In order to test whether there were differences between the recall of positive versus negative carbon labels on a brand level, another repeated measures test was performed. In this test, recall of positive and negative carbon labels on a brand level were the dependent variables. Condition and NEP were the predictors, and perceptions of influence and social norms were taken as covariates. The results show us the following:

Within subjects

There was no significant difference between the recall of positive versus negative carbon labels on a brand level ($F(1,121) = 1.578, p = .211$). Recall of positive carbon labels on a brand level ($M = .42, SD = .50$) versus a negative carbon label on a brand level ($M = .51, SD = .50$) did not significantly differed from each other.

Condition had no significant main effect on the recall of positive versus negative carbon labels on a brand level, $F(3,121) = .727, p = .538$. This shows that the recall between respondents not differed per condition. This is in contrast with both hypotheses 4 and 5.

NEP had no significant main effect on the recall of positive versus negative carbon labels on a brand level, $F(1,121) = 1.253, p = .265$. This shows that the recall between respondents with a high or low NEP did not significantly differed from each other. Participants with a low NEP recalled a positive carbon label on a brand level a bit better ($M = .38, SD = .50$) as compared to participants with a high NEP ($M = .46, SD = .50$). This effect though is not significant. Neither for a negative carbon label on a brand level did recall of low NEP respondents ($M = .50, SD = .50$) differed from high NEP respondents ($M = .52, SD = .50$). Yet it is interesting to see that negative carbon labels on a brand level are better remembers by low NEP participants as compared to high NEP participants.

No significant effect of the controlling covariate perceptions of influence ($p = .259$) was found.

Social norms appeared to correlate with recall, $F(1,121) = 4.832, p = .030$. This influence is in a positive direction for the positive carbon label on a brand level ($B = .17, p = .056$). For a negative carbon label on a brand level, this effect is in a negative direction ($B = -.02, p = .873$).

Between subjects

Regarding the different labels, there was a significant simple main effect of condition for recall, $F(3,121) = 62.372, p = .071$. This shows that there was a difference in recall between the four different labels (checkmark, traffic light, black score and colored score). Bonferroni corrected post hoc test shows that recall of a checkmark label ($M = .14, SE = .05$) differed significantly from a black score label ($M = .81, SE = .05$). On average, participants remembered a checkmark label better as compared to the black score

(*Mean difference* = .67, $p = .000$). Furthermore, recall of the traffic light label ($M = .11$, $SE = .05$) differed significantly from the black score label ($M = .81$, $SE = .05$) and the colored score ($M = .83$, $SE = .05$). This differences were on average .70 ($p = .000$) and .72 ($p = .000$) respectively. This shows that the traffic light label was better remembered as compared to the black and the colored score. No significant difference was found between the recall of black score label and the colored score label ($p = 1.000$). These findings support hypothesis 4, but are in contrast with hypothesis 5.

NEP did not had a significant simple main effect between subjects $F(1,121) = .003$, $p = .959$. This effect tells us that participants' NEP had no effect on their recall of the carbon labels on a brand level. Bonferroni corrected post hoc tests showed that the recall of labels on a brand level did not significantly differed between respondents who scored high ($M = .47$, $SE = .04$) versus low ($M = .47$, $SE = .04$) on NEP.

Social norms appeared to correlate with the recall of carbon labels on a brand level , $F(1,121) = 3.311$, $p = .071$. This influence is in a positive direction for the positive carbon label on a brand level ($B = .17$, $p = .056$). For a negative carbon label on a brand level, this effect is in a negative direction ($B = -.02$, $p = .873$).

No significant effect of the controlling covariate perceptions of influence ($p = .993$) was found.

Finally, there was a significant simple main effect of recall of positive labels on a brand level ($M = .43$, $SE = .03$) versus negative labels on a brand level ($M = .52$, $SE = .03$). On average, participants did remember a positive carbon label on a brand level better as compared to a negative label on a brand level (*Mean difference* = .09, $p = .022$). This was after the scores for recall are corrected for social norms and perceptions of influence.

Recall approximately positive versus negative label, Brand level

In this test, the approximate recall of positive versus negative carbon labels on a brand level were the dependent variables. Condition and NEP were the predictors, and perceptions of influence and social norms were taken as covariates. The results of the

test between the positive versus negative carbon labels on a brand level show us the following:

Within subjects

There was a significant difference between the recall of category of positive versus negative carbon labels on a brand level ($F(1,121) = 10.351, p = .002$). This effect shows that, if we ignore the fact whether the recall test was performed for different types of labels, the recall of the correct category between positive and negative carbon labels on a brand level significantly differed from each other. Positive carbon labels on a brand level were better ($M = .08, SD = .28$) remembered as compared to negative carbon labels on a brand level ($M = .37, SD = .49$).

Condition appeared to have a significant main effect on the recall of category of positive versus negative carbon labels on a brand level, $F(3,121) = 70.742, p = .00$. This shows that the recall between positive versus negative carbon labels differed per condition.

NEP had no significant main effect on the recall of category of the positive versus negative carbon labels on a brand level, $F(1,121) = .111, p = .740$. This shows that the recall of category between respondents with a high or low NEP did not significantly differ from each other. Participants with a low NEP recalled the category of a positive carbon label on a brand level a bit less ($M = .12, SD = .33$) as compared to participants with a high NEP ($M = .05, SD = .21$). This effect though is not significant. Neither for a negative label on a brand level did recall of category for low NEP respondents ($M = .35, SD = .48$) differ from high NEP respondents ($M = .40, SD = .49$). Still it is interesting to see that in this case, positive carbon labels are remembered a bit better by high NEP participants, while negative carbon labels are remembered better by low NEP participants.

No significant effects of the controlling covariates perceptions of influence ($p = .723$) and social norms ($p = .133$) were found.

Between subjects

Regarding the different labels, a significant simple main effect of condition for recall was found, $F(3,121) = 15.677, p = .000$. This effect tells us that there was a difference in recall of category between the four different labels (checkmark, traffic light, black score and colored score). These differences were between the checkmark label ($M = .14, SE =$

.05) and the black score label ($M = .50$, $SE = .05$). Participants remembered the checkmark label better as compared to the black score label (*Mean difference* = .36, $p = .000$). Furthermore, recall of the traffic light label ($M = .11$, $SE = .04$) differed significantly from the black score label ($M = .50$, $SE = .05$) with an average difference of .39 ($p = .000$). The colored score label ($M = .16$, $SE = .05$) was significantly better remembered as compared to the black score label ($M = .50$, $SE = .05$) (*Mean difference* = .34, $p = .000$). This shows that the black score label was less remembered as compared to all the other labels. These findings partly support hypothesis 4, and fully support hypothesis 5.

NEP did not have a significant simple main effect between subjects $F(1,121) = 1.390$, $p = .241$. This effect tells us that participants' NEP had no effect on their recall of categories. Bonferroni corrected post hoc tests showed that the carbon label recall of categories between respondents who scored high ($M = .20$, $SE = .03$) versus low ($M = .26$, $SE = .03$) on NEP did not significantly differ.

Social norms appeared to correlate with the recall of category of carbon labels on a brand level, $F(1,121) = 2.832$, $p = .095$. For a positive carbon label on a brand level, this effect is in a positive direction ($B = .20$, $p = .023$). For a negative carbon label on a brand level, this effect is also in a positive direction ($B = .07$, $p = .414$).

No significant effect of the controlling covariate perceptions of influence ($p = .822$) was found.

Finally, there was a significant simple main effect of recall of category of positive carbon labels on a brand level ($M = .08$, $SE = .02$) versus a negative carbon labels on a brand level ($M = .37$, $SE = .03$). This shows again that a positive label on a brand level was better remembered as compared to a negative label on a brand level (*Mean difference* = .29, $p = .000$).

Recall exact positive versus negative label, Product level

In order to test whether there were differences between the recall of positive versus negative carbon labels on a product level, another repeated measures test was performed. In this test, recall of positive and negative carbon labels on a product level were the dependent variables. Condition and NEP were the predictors, and perceptions

of influence and social norms were taken as covariates. The results shows us the following:

Within subjects

There was a significant difference between the recall of positive versus negative carbon labels on a product level ($F(1,121) = 2.891, p = .092$). Recall of positive carbon labels on a product level ($M = .47, SD = .50$) versus a negative carbon label on a product level ($M = .47, SD = .50$) did significantly differ from each other.

Condition had a significant main effect on the recall of positive versus negative carbon labels on a product level, $F(3,121) = 6.022, p = .001$. This shows that the recall between respondents differed per condition.

NEP had no significant main effect on the recall of positive versus negative carbon labels on a product level, $F(1,121) = .361, p = .549$. This shows that the recall between respondents with a low of high NEP did not significantly differed from each other. Participants with a low NEP recalled a positive carbon label on a product level a bit less ($M = .50, SD = .50$) as compared to participants with a high NEP ($M = .49, SD = .50$). This effect though is not significant. Neither for a negative carbon label on a product level did low NEP ($M = .46, SD = .50$) differed significantly from high NEP respondents ($M = .54, SD = .50$). Yet it is interesting to see that for labels on a product level, low NEP respondents recall both the positive and the negative label equally or better as compared to the high NEP participants.

No significant effect of the controlling covariate perceptions of influence ($p = .716$) was found.

Social norms appeared to correlate with recall, $F(1,121) = 3.696, p = .073$. For a positive carbon label on a product level, this direction is negative ($B = -.10, p = .242$). For a negative label on a product level, this effect is in a positive direction ($B = .03, p = .770$).

Between subjects

Regarding the different labels, there was a significant simple main effect of condition for recall, $F(3,121) = 55.78, p = .000$. This shows that there was a difference in recall between the four different labels (checkmark, traffic light, black score and colored score). Bonferroni corrected post hoc test shows that recall of a checkmark label ($M = .25, SE =$

.05) significantly differed from recall of the black score label ($M = .76$, $SE = .05$). Participants remembered a checkmark label better as compared to the black score label (*Mean difference* = .51, $p = .000$). The checkmark label ($M = .25$, $SE = .05$) was also significantly better remembered as compared to a colored score label ($M = .90$, $SE = .05$) (*Mean difference* = .65, $p = .000$). Furthermore, recall of the traffic light label ($M = .10$, $SE = .05$) differed significantly from the black score label ($M = .76$, $SE = .05$) and the colored score ($M = .90$, $SE = .05$). This mean differences were on respectively .66 ($p = .000$) and .80 ($p = .000$). This shows that the traffic light label was better remembered as compared to the black score label and the colored score label. No significant difference was found for the recall between the black score label and the colored score label ($p = .286$). Interesting to see is that the black score label was better remembered as compared to the colored score. However, they did not differed significantly from each other. These findings support hypothesis 4, and contradict with hypothesis 5.

NEP did not had a significant simple main effect between subjects $F(1,121) = .283$, $p = .596$. This effect tells us that participants' NEP had no effect on their recall of the carbon labels on a product level. Bonferroni corrected post hoc tests showed that the recall of labels on a product level did not significantly differed between respondents who scored high ($M = .52$, $SE = .04$) versus low ($M = .49$, $SE = .04$) on NEP.

Perceptions of influence appeared to correlate with the recall of carbon labels on a product level, $F(1,121) = 5.984$, $p = .016$. For a positive carbon label on a product level, this effect was in a negative direction ($B = -.16$, $p = .072$). For a negative label on a product level, this effect was also in a negative direction ($B = -.12$, $p = .182$).

No significant effect of the controlling covariate social norms ($p = .914$) was found.

Finally, there was no significant simple main effect of recall of positive carbon labels on a product level ($M = .51$, $SE = .04$) versus negative carbon labels on a product level ($M = .50$, $SE = .03$) after correcting the values for social norms and perceptions of influence. On average, participants did remember a negative carbon label on a product level a little bit better as compared to a positive carbon label on a product level (*Mean difference* = .01). However, this difference is very small and not significant ($p = .828$).

Recall approximately positive versus negative label, Product level

In this test, the approximate recall of positive versus negative carbon labels on a product level were the dependent variables. Condition and NEP were the predictors, and perceptions of influence and social norms were taken as covariates. The results of the test show us the following:

Within subjects

There was no significant difference between the recall of category of positive versus negative carbon labels on a product level ($F(1,121) = 1.236, p = .268$). Recall of positive carbon labels ($M = .21, SD = .41$) versus negative carbon labels ($M = .35, SD = .48$) did not significantly differed from each other.

Condition had a significant main effect on the recall of category of positive versus negative carbon labels on a product level, $F(3,121) = 43.209, p = .000$. This shows that the recall of category between positive versus negative carbon labels on a product level differed per condition.

Nep had no significant main effect on the recall of category of the positive versus negative carbon labels on a product level, $F(1,121) = .798, p = .373$. This shows that the recall of category between respondents with a high or low NEP did not significantly differed from each other. Participants with a low NEP recalled the category of a positive carbon label on a product level a bit better ($M = .20, SD = .40$) as compared to participants with a high NEP ($M = .23, SD = .43$). This effect though is not significant. Neither for a negative carbon label on a product level did low NEP ($M = .30, SD = .46$) differed significantly with high NEP respondents ($M = .40, SD = .49$). Still it is interesting to see that for carbon labels on a product level, low NEP respondents recall the category of both the positive and the negative carbon label better as compared to high NEP participants.

No significant effects of the controlling covariates perceptions of influence ($p = .387$) and social norms ($p = .107$) were found.

Between subjects

Regarding the different labels, a significant simple main effect of condition for recall of category was found, $F(3,121) = 14.310$, $p = .000$. This effect tells us that there was a difference in recall of category between the four different labels (checkmark, traffic light, black score and colored score). Bonferroni corrected post hoc test shows that these differences were between the checkmark label ($M = .24$, $SE = .05$) and the black score label ($M = .56$, $SE = .05$). Participants remembered the checkmark label better as compared to the black score (*Mean difference* = $.32$, $p = .000$). The checkmark label was almost equally well remembered as the colored score labels in this situation. Furthermore, recall of category of the traffic light label ($M = .10$, $SE = .05$) differed significantly from the black score label ($M = .56$, $SE = .05$), with a mean difference of $.46$ ($p = .000$). Black score label ($M = .56$, $SE = .05$) also differed significantly from colored score label ($M = .24$, $SE = .05$). Colored scores were with a mean difference of $.32$ better remembered compared to the black scores ($p = .000$). This shows again that all the carbon labels were better remembered as compared to the black score label. This is in line with hypothesis 4 and 5.

NEP did not had a significant simple main effect between subjects $F(1,121) = .431$, $p = .513$. This effect tells us that participants' NEP had no effect on their recall of categories. Bonferroni corrected post hoc tests showed that the carbon label recall of categories between respondents who scored high ($M = .30$, $SE = .04$) versus low ($M = .27$, $SE = .04$) on NEP did not significantly differ.

No significant effects of the controlling covariates perceptions of influence ($p = .508$) and social norms ($p = .571$) were found.

Finally, there was a significant simple main effect of recall of category of the positive carbon labels on a product level ($M = .22$, $SE = .04$) versus the negative carbon labels on a product level ($M = .35$, $SE = .03$). This shows that a positive label on a product level was better remembered as compared to a negative label on a product level (*Mean difference* = $.13$, $p = .001$), after correcting the scores for social norms and perceptions of influence.

Conclusion recall

Previous section thoroughly displayed the results regarding recall. The following conclusions can be drawn:

H4: The checkmark and the traffic light labels are better recalled as compared to a black score and a colored score label.

It was found that both the checkmark and the traffic light label were significantly better remembered compared to the black score labels and the colored score labels. If a label was negative, a checkmark and a traffic light label were better remembered compared to the black and colored score labels. As a result, **hypothesis 4** will be **accepted**.

H5: A colored score label will be recalled better as compared to a black score label.

Mixed results regarding hypothesis 5 are shown in the previous section. If label information is on a product level, there is a significant difference regarding the recall of a black or colored score label. Participants made more errors in recalling the exact score of the colored label, compared to the black score. This means that the black score was significantly better remembered. However, if the information is on a brand level, there was no difference in recall between the black versus the colored score. This is in contrast with **hypothesis 5**, which as a result will be **rejected**.

Taken this together, it can be stated that hypothesis 4 is accepted, and that hypothesis 5 is rejected.

Assessment label test

Results of the label assessment (appendix XIII) show that the four different labels (checkmark, traffic light, black score and colored score) are ranked on average about the same. A significant effect of type of label was found for “*The carbon label provided me with enough information to estimate the carbon emission friendliness of the product*” ($F(3,130) = 2.850, p = .040$). Bonferroni corrected post-hoc test showed that this significant difference occurred between the checkmark label and the colored score label. This implies that the participants in the colored score condition significantly more often stated that the label gave them enough information to estimate the carbon friendliness of the product they had to “buy” as compared to the participants who answered the questions for the checkmark label condition.

Even though no further significant differences between the four labels were found, it is interesting to see that while checkmark labels and traffic light labels were the best remembered as well the best liked carbon labels, participants indicated that they understood the black score label ($M = 5.27, SD = 1.40$) and colored score ($M = 5.22, SD = 1.16$) labels better as compared to the checkmark label ($M = 5.00, SD = 1.41$). Traffic light label still scored higher on this question ($M = 5.71, SD = .87$), and thus was indicated as “best understandable”. This shows that there is no clear evidence that the checkmark and traffic light labels are indicated as better understandable as compared to the black- and colored score labels. This is in contrast with **hypothesis 1**, which as a result will be **rejected**.

Furthermore, ratings regarding the question “*The carbon label provided me with enough information to estimate the carbon emission friendliness of the product*” are higher for both black score label ($M = 4.55, SD = 1.50$) and colored score label ($M = 5.09, SD = 1.23$) as compared to the checkmark label ($M = 4.06, SD = 1.78$). Colored score label ($M = 5.09, SD = 1.23$) scored even higher than traffic light ($M = 4.79, SD = 1.13$) on this question. Even though the differences in ratings were not significant, it is interesting to see that both the “score labels” were indicated as providing more information, compared to the checkmark and traffic light labels.

Discussion

This report was written in order to provide an insight in how different types of label influence consumer preferences. Towards these labels, a distinction was made between information on a brand level versus on a product level. This chapter will draw conclusions based upon the results by summarizing the results (based on the hypotheses) and by answering the (sub)research questions. Hereafter, limitations, a general conclusion and managerial implications will be discussed.

Willingness to pay (WTP)

Current research found that information on the carbon label on a brand level versus a product level had a significant effect on respondent's WTP, when the carbon label was positive. This effect did not appear for negative carbon labels. However, if the mean scores for WTP for fridges with carbon labels on a brand level versus on a product level were corrected for perceptions of influence and social norms, significant differences appeared between the WTP for fridges with a negative carbon label on a brand level versus a product level. For both the positive and negative carbon labels, participants were willing to pay more for a fridge with a carbon label on a product level as compared to a fridge with a carbon label on a brand level.

A possible explanation for this finding is the fact that current research used unknown brand names for the manipulation of brand level versus product level. It is possible that this diminished the effect of brand knowledge. However, this was a well thought through choice, in order to overcome brand preferences. Research has shown that whenever a buyer has had a previous experience with a brand, this information is likely to be a dominant factor in the decision making process (Monroe, 1976). In order to overcome this bias, participants in current research saw the label on a certain level, followed by a short text. Subsequently, a manipulation check was performed in order to see whether the participants understood that the label was on a brand level or on a product level. Participants who incorrectly answered this question were completely excluded from the research and thus we only used results of participants who understood that the label they saw was on a brand level or on a product level. Till now, little research has been conducted on the use of unknown brands. Monroe (1976) used both known and unknown brands in order to measure the effect of brand familiarity on consumer preferences. An interesting finding of his research was that if a brand was unfamiliar to consumers, they made more use of cognitive factors. Linked to current research, this might imply that, as a result of brand unfamiliarity, participants made more cognitive

decisions while they were answering questions about the unknown brands and thus have given well evaluated answers.

Liking and attractiveness

Regarding positive carbon labels, a significant difference between the rates of liking for a fridge with a carbon label on a brand level versus a fridge with a label on a product level was found. Participants liked a fridge with a positive carbon label on a product level more compared to a fridge with a positive label on a brand level.

A negative carbon labels appeared not to have an effect on respondents' indicated rates of liking when a distinction was made between brand level and product level. With regard to attractiveness, significant higher rates of attractiveness were measured for fridges with a positive carbon label on a product level compared to fridges with a positive carbon label on a brand level. Negative carbon labels seem to score the same on attractiveness for brand level versus product level. Yet is it interesting to see that negative labels on a brand level are rated as less attractive compared to products with a negative label on a product level. This is in line with our other findings in the results section, which showed a trend that a product level is preferred over a brand level.

A significant difference between ratings of liking for positive carbon labels versus negative carbon labels was found. Positive carbon labels were liked more compared to negative carbon labels. Also regarding attractiveness, positive carbon labels are indicated as more attractive compared to negative carbon labels. Interesting is the fact that there was no significant difference in ratings of attractiveness between fridges with a positive versus a negative label on a brand level. However, when this test was repeated with corrected values for social norms and perceptions of influence, this effect became significant. Hence, a negative label with carbon information on a brand level was rated as less attractive compared to a positive label with information on a brand level.

This finding might seem logically, which could be the result of unconscious processing of the different label elements, which thus should be mentioned. For instance, green and red might have a large unconscious influence on consumers, because they instantly deem such a color as good or bad. Of course we used this distinction in order to frame a product as "good/environmental friendly" or "bad/environmental unfriendly", but yet this is making it difficult to only ascribe the results to carbon emission interpretations. Moreover, packaging of consumer goods influence consumer behavior and are riddled with warnings and other information. The current study only used the carbon label without any other labels, making it less easy to generalize to real world application, where it would

have to compete with lots of other information and labels. In current study, there is full focus on the carbon label and choice overload is nearly impossible. Further research should be done in order to assess how much a consumer cares about the carbon emissions compared to other information and how much it influences consumer decision-making when the carbon label is combined with other information on a product of package.

Recall

Looking at recall of the positive labels, it is interesting to see that while carbon labels on a brand level were rated as less favorable, the recall of positive carbon labels on a brand level was significantly higher compared to positive carbon labels on a product level. Perhaps even more interesting is the fact that this effect does not exist for negative carbon labels; no difference was found between the exact nor the category recall of the negative carbon labels whether they were on a brand level or on a product level.

At first sight, there seemed to be no significant difference between the recall of positive versus negative carbon labels on a brand level. If the scores of recall were corrected for social norms and perceptions of influence, a significant effect appeared. In this case, participants did remember a positive carbon label on a brand level better compared to a negative label on a brand level. The same applied to positive versus negative carbon labels on a product level. Meaning that after correcting the scores for social norms and perceptions of influence, there appeared to be a significant difference in recall, which showed that a positive label on a product level was better remembered compared to a negative label on a product level. Interesting here is that negative carbon labels on a product level were slightly better remembered as compared to negative carbon labels on a brand level. This was opposed to the findings regarding positive carbon labels, where carbon labels on a brand level were better remembered. Even though this effect though was not significant, a visible trend was observed with regard to these findings.

NEP value

NEP appeared to have an effect on WTP for fridges with positive carbon labels. Participants with a high NEP indicated a higher WTP for fridges with a carbon label on both brand and product level, compared to participants with a low NEP value. High NEP participants were willing to pay the most for a fridge with a positive carbon label on a product level. NEP had no effect on the WTP regarding negative carbon labels, and thus

high and low NEP participants were willing to pay on average the same amount of money for the fridges with a negative carbon label.

An interesting finding regarding participants' WTP was the fact that for carbon labels on a product level, low NEP participants were willing to pay more for a negative label as compared to participants with a high NEP. This is opposed to the appearing trend that high NEP participants have a higher WTP as compared to low NEP participants. NEP seems to have a strong effect on the WTP as a whole. Participants who scored high on NEP gave in general higher WTP's towards fridges with a positive label, while a negative label seemed to devaluate the fridge more, since lower WTP's were given by high NEP participants. The opposite is true for participants who scored low on NEP.

NEP is a scale which has been investigated thoroughly. Current research used the 15-item version of the NEP scale. This scale was proven to be more reliable than the existing 5, 7, 8 or 10 items scales, but less reliable as compared to the 6-item scale (Hawcroft and Milfont, 2010). It was believed that the 6-item scale did not cover all our questions. In order to make the scale as reliable as possible, recommendations of Hawcroft and Milfont (2010) were followed and thus current research used the revised and improved scale, and a 7-point Likert scale for the answers instead of the original 5-point scale.

Large numbers of recent studies sum up all items into a single measure of environmental attitude, and so did we. This score was called NEP value and based on the median split method, a distinction was made between high and low NEP respondents. Summing up all items into a single measure might result in information losses (e.g., Dunlap et al., 2000). However, existing reviews support this approach and therefore this should not be a problem regarding validity in current study (Hawcroft and Milfont, 2010).

High NEP participants indicated higher rates of liking and attractiveness for fridges with a positive carbon label on a product level as compared to fridges with a positive carbon label on a brand level. Interesting here is that ratings of liking and attractiveness for fridges with negative carbon labels do not differ between low and high NEP respondents. For fridges with a label on a product level, NEP appeared to have a significant effect on liking and attractiveness. Participants with a low NEP gave lower ratings of liking and attractiveness to a fridge with a positive carbon label on a product level compared to participants with a high NEP. Interestingly, low NEP participants gave higher ratings of liking towards fridges with negative carbon labels on a product level as compared to high

NEP participants. NEP had no significant influence on ratings of attractiveness for negative carbon labels, but yet, negative carbon labels on a brand level were liked more compared to positive labels on a brand level by low NEP participants compared to high NEP participants. Even though this effect was not significant, the appearing trend might be interesting. For now, we can state that high NEP participants prefer environmental friendly products, but it cannot be stated that low NEP participants make no distinction between environmental friendly and unfriendly products.

Regarding this finding, it is important to take into account that the answers with regard to environmental concerns are self-reported. Self-reported behavior is vulnerable for biases towards “ecological correctness” (Kaiser et al., 2003). This might imply that socially desirable answers are given and that participants’ scores of for example NEP value are unrealistically high. However, even if this is the case, this does not harm our results, since current research measured differences in willingness to pay and in ratings of liking and attractiveness. Yet, these results are based on whether a person scored high or low on NEP value, beliefs in social norms, and perceptions of influence. In order to correct these values for socially desirable answers, the high and low values were based on median split scores on these variables.

Even though no interaction effects of NEP on type of label were found for liking and attractiveness, it is interesting to see that high NEP participants rate the positive checkmark and traffic light label higher compared to the positive black and colored score labels in terms of liking. Rates of attractiveness are almost equal to ratings of low NEP participants. Furthermore, it can be seen that low NEP participants make a distinction between environmental friendly and unfriendly fridges, as opposed to our hypothesis. Further research in this area is required.

As a final result regarding NEP, it appeared that NEP had nearly no effect on recall. This might be an interesting finding, since you might expect that high NEP subjects would perform better on remembering the labels. While no statistical significant effects of NEP were found, there was a remarkable trend which showed that low NEP respondents tend to better recall both the positive and negative carbon labels on a brand level and on a product level, compared to high NEP participants. Furthermore, when positive versus negative carbon labels on a brand level were compared, low NEP participants scored slightly higher on recall compared to high NEP participants. When brand versus product level was compared, low NEP participants also remembered more labels as compared to

high NEP participants. This suggests that, even though not significant, low NEP participants recall the labels better, compared to high NEP participants. This might be interesting for future research.

Type of label (check mark, traffic light, black score or colored score)

No effect of type of label was found for neither positive nor negative carbon labels with regard to WTP. This implies that the type of label (check mark, traffic light, black score or colored score) does not affect participants' WTP, meaning that participants were willing to pay about the same amount of money among the different labels.

The type of label neither seemed to have an effect on ratings of attractiveness for positive framed labels. Ratings of attractiveness among the checkmark, traffic light, black score and colored score labels were about the same, if the label was positive. Regarding negative labels, significantly higher ratings were given to a black score label compared to a traffic light label and to a colored score label. This finding was partly in line with hypothesis 3. However, only if the carbon labels were negative, a black score was indicated as "most attractive" type of label. When we compare these findings with results regarding "liking", it can be seen that checkmark and traffic light were indicated as "most likable" labels (when a carbon label was positive). Significant differences in ratings of liking for positive labels were found between the checkmark and the black score, and between the traffic light and the black score. As a result, it can be said that the black score was the "most unlikable" from all four labels, if a label was positive.

For negative carbon labels, condition also had a significant effect on ratings of liking. An interesting finding here was that the black score was liked more compared to the traffic light label. However, it can be stated that the checkmark and the traffic light labels are more "liked" and seen as "more attractive" as compared to the black score and colored score label. As a result, we cannot conclude that a score label (black or color) is preferred over a checkmark and a traffic light label. Yet, there is strong evidence supporting that a score label is seen as attractive when a carbon label is negative.

Upham and colleagues (2011) state a reasonable explanation for this finding. They show that participants in their research understand that a high score differs from a low score, but that this is not sufficient to allow them to engage meaningfully with this information. As a result, they state that labels containing a comparing element (like a traffic light; which has three colours) give more meaning towards the participants. This might be the cause of preference for labels with a reference point (checkmark and traffic light) compared to a plain score. Adding colors to the score label was meant to give

participants a reference point, but this did not lead to more preference of the label compared to a black score label.

Regarding the positive carbon labels on a brand level or product level, the checkmark and the traffic light label were both significantly better remembered compared to the black score labels and the colored score labels. If a label was negative, a checkmark and a traffic light label were better remembered compared to the black and colored score labels. The black score label was the worst remembered label in all cases.

When information on a carbon label was on a brand level, the checkmark label and the traffic light label were again better remembered compared to the black score label. Traffic light label also differed significantly in recall scores from colored score labels. When the scores for recall are corrected to recall of category, the colored score label was also better remembered compared to the black label. When information on a carbon label was on a product level, again the checkmark label and the traffic light label were better remembered compared to both the black score label and the colored score label. For negative carbon labels, black score labels were slightly better remembered compared to colored score labels, which is not in line with previous findings. However, this difference was not significant. Furthermore, as opposed to our hypothesis, a colored score label was not recalled better than a black score label.

Regarding the four types of labels, the colored score label was indicated as significantly providing more information in order to estimate the carbon friendliness of a product compared to the checkmark label. Furthermore, no significant differences were found. However, a trend was observed, showing that participants indicated that the traffic light score and the colored score gave them the most important information. There was no significant difference between the checkmark and the black score labels. Traffic light scored the highest on almost all of the questions regarding the information about the four types of labels, and thus can be seen as the overall winner which participants prefer. However, there is no clear evidence that the checkmark and traffic light labels are indicated as better understandable as compared to the black- and colored score labels.

A possible explanation for this finding is the fact that current research used existing labels in a new context. The checkmark and traffic light labels are known labels, but not in the area of carbon labels. A consistent mechanism, communicating the environmental impact of a product or brand, would convey a meaningful message (Cohen and Viscusi,

2012), leading to better understanding. However, we are dealing with a new topic in current research, still in an explanatory stage, trying to find out which label should be used in order to communicate a clear and uniform message to the audience about environmental friendliness and impact. As Upham and colleagues (2011) stated, it takes time and general guidelines to set a certain norm with regard to labels. An example could be a label showing Guideline Daily Amounts (GDA's) on a product. Many people know what terms on these kinds of labels mean, and thus they are able to easily convert the message into useful information. Furthermore, multiple researchers suggest that additional information on the label is important and necessary (Cohen and Vandenberg, 2012; Upham, Dendler and Bleda, 2011), in order to clarify what the label is saying and making the label easier to interpret. Current research only mentioned whether a fridge with a label was carbon friendly or unfriendly. A recommendation for future research is to provide more information next to the carbon label. Furthermore, the knowledge regarding carbon emissions of general public could be rather low. This would make the presented information difficult to interpret. Based on this, we recommend to provide a brief but clear description about carbon emissions and its effects in future research. Current research did not provide this description, which might have influenced results in a way that participants did not completely understand the different labels.

It is noteworthy that current research used the .1 significance level, instead of the commonly used .05 level. This has been an elaborated choice, since this research was in such an explorative state that very little findings appeared. However, many results seemed to be nearly significant, and thus suggested a pattern. It might be interesting to conduct further research regarding the findings which have a higher significance level than 0.05, but lower than 0.1, in order to elaborate the findings more deeply.

Finally, it is important to emphasize that current research did not look at a real world situation where people had to make real choices, costing them real money. It might be the case that participants find one type of label the most easy to interpret, but they do not use it at all in their decision making. Another type of label is not preferred but is remembered at the point of decision making which influences in some way the end result. This means we deal with different aspects, being consumer preferences regarding the different types of labels, and the influence of the types of labels. Interesting would be to use the findings of current study, in order to conduct a research in which participants can actually purchase a product, in order to see how these findings apply in the real world.

Limitations

Survey duration

After a small pre-test of the survey which was used for current research, there were some complaints about the duration of the survey. Hereafter, some adjustments were made in order to reduce the duration. However, the survey would be used for two researches. In order to reward respondents for their participation in the long survey, they were promised to become participant in a lottery of two iPad's after filling in the survey. By this, it was aimed to obtain serious results and no hurried completion of the survey. Yet we found out during the collection of the data that quite a lot of respondents quitted the survey before completing it. This has led to a loss of data. We also removed some distrustful data like a respondent who gave the same answer 15 times in a row, or respondents who finished the survey remarkably fast, since it was believed that some respondents only filled in the survey in order to win the iPad. We are aware that there is a trade-off between having brief but less accurate measures, and having long but accurate ones. Therefore we decided to proceed with the long survey.

Aspects that further complicate the things is the fact that the products which are used in this study (fridges) are often products that are not bought on an impulse, but are rather time and energy consuming searches before the actual purchase is made. This might imply that emotions play less a role as compared to products like food and clothes which are more sensitive for impulsive buying (Rook et al., 1995). As a result, ratings for WTP might be more thought through by respondents compared to ratings for liking and attractiveness. Yet we do not think this has influenced our results. Moreover, after launching the survey the participants ran into the so called "defensive denial hypothesis" (Tyler et al., 1982). This hypothesis states that, under high-cost conditions, individuals' attitudes become less important, compared to low-cost situations. Current study dealt with a high-cost condition, which is purchasing a fridge. Based on the hypothesis of Tyler et al. (1982), it can be assumed that environmental attitudes of respondents were downgraded in the situation of purchasing the fridge. However, current study started with questions about the environment, beliefs in climate change and beliefs about influencing climate change. By doing this, it was aimed to trigger participants' environmental attitude. Besides that, current study used mean scores, so we are convinced that overall scores will be accurate and thus the use of high-cost situations will not have influenced the validity of current study.

Furthermore, this study thoroughly analyzed all the results, using different analyses. The levels of product and brand were split and the same applied to positive and negative labels. Even though this implies that the results are very well underpinned, this has led to rejection of almost all the hypotheses. This is not a bad thing per se, however, many analyses supported a part of a certain hypothesis. This means that, while the recommendations in this report follow directly from empirical results, caution should be exercised in interpretation.

Current research used different operationalizations of the variable “preference”. This leads to different results. Variations in the way we measured the construct influenced whether or not we confirm our hypotheses. For example, we measured liking and attractiveness. Results regarding both of these constructs were used in order to state results regarding the hypotheses with the variable “preference”.

This links to one of the biggest limitations of current research, being the fact that the survey used was not perfectly built based on the hypotheses. As a result, not all hypotheses were perfectly measured. This could have been overcome by performing a first analysis, at the moment the first results came in. This was not done. As a result, it was not possible to analyze two of the hypotheses regarding the covariates social norms, and perceptions of influence (*H13, H14*). This is a pity, since these covariates appear to have a moderating effect (either reinforcing or weakening) on the relationship between the conditions and rates of liking, attractiveness, and WTP. This should be interesting to take into account in future research.

Another limitation can be found in the fact that current study analyzed positive and negative carbon labels on appliances (fridges). It is important to realize that findings of current research can be very different for fast moving consumer goods, and thus generalization should be handled carefully.

The same applies for a participant who relies his rating more on demographic variables versus a participant who gives higher weight towards product-attributes. For future research, this implies that the direction of influence of the mentioned determinants should be taken into account.

Conclusion

Current study examines the effect of brand level information versus product level information on consumer preferences. Participants are willing to pay more for a fridge with a carbon label on a product level, regardless of the carbon label being positive or negative. With regard to liking and attractiveness, participants were more liking and more attracted to a positive carbon label on a product level compared to a positive carbon label on a brand level. Negative carbon labels on a product level and on a brand level are equally “liked” and seen as equally “attractive”.

A high NEP value results in a higher WTP for positive carbon labels on a product level as well as on a brand level. A high NEP also leads to higher rates of liking and attractiveness for positive carbon labels compared to negative carbon labels. There is no effect of NEP on WTP, liking, and attractiveness for fridges with a negative carbon label. Type of label (checkmark, traffic light, black score and colored score) does not influence participants' WTP and rates of attractiveness. However, rates of liking for positive carbon labels are influenced by the type of label; the checkmark and the traffic light label are significantly liked more compared to the black- and colored score labels. No significant differences in rates of liking are found for negative carbon labels. However, negative carbon labels differ significantly in rates of attractiveness in a way that a black score label is seen as more attractive compared to a traffic light and a colored score.

Even though carbon labels on a brand level are rated as less favorable compared to carbon labels on a product level, positive carbon labels on a brand level are significantly better remembered compared to positive carbon labels on a product level. No difference in recall is found for negative carbon labels on a product level versus on a brand level. Positive carbon labels on a brand level are better remembered compared to negative carbon labels on a brand level. If the given information is on a product level, no difference in recall exists between positive versus negative labels. Regarding positive carbon labels, the checkmark and the traffic light label are both significantly better remembered compared to the black score labels and the colored score labels. Also, a negative checkmark is better remembered compared to other negative types of labels.

Long story short, it can be stated that a label on a product level is the overall winner, if the label is positive. If the label is negative, it does not really matter whether the information is on a brand level or on a product level. Yet it is recommendable to make use of a black score label if the carbon information is negative. If the carbon information is positive, it is advised to use a checkmark or a traffic light label.

Recommendations for future research

Taken this together, it can be said that carbon labelling is quite a complex matter. But yet, there is a lot of potential for carbon labels. Even though some researchers state that it is unlikely that lots of consumers will use carbon labels as a basis for their product selection (Upham et al., 2011; Giorgi and Fell, 2009), current research shows that positive carbon labels are liked more and seen as more attractive compared to negative carbon labels. Besides, current research showed that consumers are willing to pay more for carbon friendly products. However, most people do not think about energy consumption or environmental impact when buying appliances (Lyndhurst, 2007), buying and preparing food, or when they travel (King et al., 2009). If attitudes towards environmental issues can be changed, it is possible that consumers' decision making and purchase behavior can be influenced (Pickett-Baker and Ozaki, 2008). A possible way to gain attention towards environmental issues like carbon emission, is to set up general guidelines. Currently, there are no international guidelines regarding carbon labelling (Upham et al., 2011). As a result, companies are quite free in their decision-making regarding carbon labels, which leads to different labels in stores and on products. This leads to confusion among consumers and therefore a general understanding of the labels is lacking. As mentioned before, a uniform label on a product category might result in a better understanding of the label and this might possibly influence consumer behavior. As with any other environmental label, carbon labels will have to compete for consumers' attention. However, current research shows the potential of carbon labels and thus this is a subject which merits much more research attention.

Current research was not able to perfectly analyze the effect of the covariates "social norms" and "perceptions of influence", while there appears to be an moderating effect of these variables on rates of liking, attractiveness and willingness to pay. Future research should be conducted with this in mind, in order to verify this moderating effect. If this effect is existing, this could be very interesting for the marketing departments of certain companies. For example, if high social norms appear to have a reinforcing effect on a consumers' liking/attractiveness/WTP for a environmental product (with a positive label), this could be used in an marketing E-mail, offering a product with a positive carbon label. If a text is added which is framed in a way that is steers how a consumer is expected to

act, it could lead to higher rates of liking/attractiveness/WTP, and thus to increased purchase intentions among consumers with high social norms.

Besides that, current study showed different results regarding positive and negative carbon labels, either on a product level or on a brand level. Like mentioned before, in this situation, the focus was fully on the one single carbon label. This is not the case in “real life”. This makes it necessary to perform more research, in order to find out whether information next to the carbon information leads to different results. In this case, the yet existing knowledge regarding carbon labels should be taken into account as well, in order to determine how much information (just a traffic light, or with a short or long description) is required.

This report showed that low NEP participants rated negative labels higher as compared to high NEP participants, when the label was on a brand level. Liking for positive labels on a brand level was ranked the same among high versus low NEP participants. This clearly states that high environmental concerned consumers prefer environmental friendly products. This could be a very interesting basis for environmental concerned companies or brands, in order to stimulate their turnover or brand knowledge by for example using carbon labels on their products. Since these companies/brands are very likely to attract high NEP value consumers, continuation of research regarding this subject could be valuable for them.

However, current research suggested that (even though not significant) low NEP participants recalled both the positive and negative carbon labels on a brand level and on a product level better, compared to high NEP participants. It is recommended to further investigate this effect, since it suggest that low NEP participants pay more attention towards the carbon label.

Literature

Algesheimer, R., Dholakia, U. M., & Herrmann, A. (2005). The social influence of brand community: Evidence from European car clubs. *Journal of marketing*, 69(3), 19-34.

Andy Field. (2009). *Discovering statistics using SPSS*. Sage publications.

Arora, R., & Stoner, C. (1996). The effect of perceived service quality and name familiarity on the service selection decision. *Journal of Services Marketing*, 10(1), 22-34.

Batra, R., & Stayman, D. M. (1990). "The Role of Mood in Advertising Effectiveness," *Journal of Consumer Research*, 17 (September), 203–214.

Biswas, A. (1992). The moderating role of brand familiarity in reference price perceptions. *Journal of Business Research*, 25(3), 251-262.

Bowlig, S., & Gibbon, P. (2009). Counting Carbon in the Marketplace: Part I—Overview Paper. *documento de trabajo no publicado*.

Burchell, K., Rettie, R., & Patel, K. (2013). Marketing social norms: Social marketing and the 'social norm approach'. *Journal of Consumer Behaviour*, 12(1), 1-9.

Cacioppo, J. T., Petty, R. E., Feinstein, J., & Jarvis, W. B. G. (1996). Dispositional differences in cognitive motivation: The life and times of individuals varying in need for cognition. *Psychological Bulletin*, 197-253.

Carbon Trust (2008), Carbon Trust Product Carbon Footprinting: The New Business Opportunity. Experience from Leading Companies, Carbon Trust, London.

Carrington, M. J., Neville, B. A., & Whitwell, G. J. (2010). Why ethical consumers don't walk their talk: Towards a framework for understanding the gap between the ethical purchase intentions and actual buying behaviour of ethically minded consumers. *Journal of Business Ethics*, 97(1), 139-158.

Cialdini, R. B., Reno, R. R., & Kallgren, C. A. (1990). A focus theory of normative conduct: recycling the concept of norms to reduce littering in public places. *Journal of personality and social psychology*, 58(6), 1015.

Cohen, M. A., & Vandenberg, M. P. (2012). The potential role of carbon labeling in a green economy. *Energy Economics*, 34, S53-S63.

Cohen, M. A., & Viscusi, W. K. (2012). The role of information disclosure in climate mitigation policy. *Climate Change Economics*, 3(04), 1250020.

Crowley, T. J. (2000). Causes of climate change over the past 1000 years. *Science*, 289(5477), 270-277.

Diekmann, A., & Preisendörfer, P. (2003). Green and greenback the behavioral effects of environmental attitudes in low-cost and high-cost situations. *Rationality and Society*, 15(4), 441-472.

Dodds, W. B., & Monroe, K. B. (1985). The effect of brand and price information on subjective product evaluations. *Advances in consumer research*, 12(1), 85-90.

Dolan, T. J., & McGeorge, K. K. (1994). Relationships between spectator identification and spectators' perceptions of influence, spectators' emotions, and competition outcome. *Journal of sport & exercise psychology*, 16, 347-364.

Dunlap, R. E., & Jones, R. E. (2003). Environmental attitudes and values. *Encyclopedia of psychological assessment*, 1, 364-369.

Dunlap, R. E., & Van Liere, K. D. (1978). The "new environmental paradigm": A proposed measuring instrument and preliminary results. *Journal of Environmental Education*, 9, 10-19.

Dunlap, R.E., Van Liere, K.D., Mertig, A.G. & Jones, R.E. (2000). New Trends in Measuring Environment Attitudes: Measuring endorsement of the New Ecological Paradigm: A revised NEP-scale. *Journal of Social Issues*, 56 (3), 425-442.

Edwards-Jones, G., Plassmann, K., York, E. H., Hounsome, B., Jones, D. L., & Milà i Canals, L. (2009). Vulnerability of exporting nations to the development of a carbon label in the United Kingdom. *environmental science & policy*, 12(4), 479-490.

Ellen, P. S., Wiener, J. L., & Cobb-Walgren, C. (1991). The role of perceived consumer effectiveness in motivating environmentally conscious behaviors. *Journal of Public Policy & Marketing*, 102-117.

Eppler, M. J., & Mengis, J. (2004). The concept of information overload: A review of literature from organization science, accounting, marketing, MIS, and related disciplines. *The information society*, 20(5), 325-344.

Geller, J. M., & Lasley, P. (1985). The new environmental paradigm scale: A re-examination. *The Journal of Environmental Education*, 17(1), 9-12.

Giorgi, S., & Fell, D. (2009). EV0402: Public Understanding of Links between Climate Change and (i) Food and (ii) Energy Use: Final Report. A report to the Department for Environment, Food and Rural Affairs.

Gössling, S., & Buckley, R. (2014). Carbon labels in tourism: Persuasive communication?. *Journal of Cleaner Production*.

Griskevicius, V., Tybur, J. M., & Van den Bergh, B. (2010). Going green to be seen: status, reputation, and conspicuous conservation. *Journal of personality and social psychology*, 98(3), 392.

Grankvist, G. and Biel, A. (2001). The Importance of Beliefs and Purchase Criteria in the Choice of Eco-Labelled Food Products. *Journal of Environmental Psychology*, 21, pp. 405-410.

Harlow, L. L. (2014). *The Essence of Multivariate Thinking: Basic Themes and Methods: Basic Themes and Methods*. Routledge.

Hawcroft, L. J., & Milfont, T. L. (2010). The use (and abuse) of the new environmental paradigm scale over the last 30 years: A meta-analysis. *Journal of Environmental psychology, 30*(2), 143-158.

Hawthorne, K. M., Moreland, K., Griffin, I. J., & Abrams, S. A. (2006). An educational program enhances food label understanding of young adolescents. *Journal of the American Dietetic Association, 106*(6), 913-916.

Hines, J. M., Hungerford, H. R., & Tomera, A. N. (1987). Analysis and synthesis of research on responsible environmental behavior: A meta-analysis. *The Journal of environmental education, 18*(2), 1-8.

Hopper, J. R., & Nielsen, J. M. (1991). Recycling as altruistic behavior normative and behavioral strategies to expand participation in a community recycling program. *Environment and behavior, 23*(2), 195-220.

Hoyer, W. D., & Brown, S. P. (1990). Effects of brand awareness on choice for a common, repeat-purchase product. *Journal of consumer research, 14*1-148.

Hoyer, W. D. (1984). An examination of consumer decision making for a common repeat purchase product. *Journal of consumer research, 8*22-829.

Hrebiniak, L. G. (1974). Effects of job level and participation on employee attitudes and perceptions of influence. *Academy of Management Journal, 17*(4), 649-662.

Hwang, M. I., & Lin, J. W. (1999). Information dimension, information overload and decision quality. *Journal of information science, 25*(3), 213-218.

Kaiser, F. G., Doka, G., Hofstetter, P., & Ranney, M. A. (2003). Ecological behavior and its environmental consequences: A life cycle assessment of a self-report measure. *Journal of Environmental Psychology, 23*(1), 11-20.

Kaiser, F. G., Wölfling, S., & Fuhrer, U. (1999). Environmental attitude and ecological behaviour. *Journal of environmental psychology, 19*(1), 1-19.

Kellogg, W. W. (1979). Effects of Human Activities on Global Climate. *WMO Tech. Note No. 156*, World Meteorological Organization, Geneva, Switzerland.

Kellogg, W. W., & Schwart, R. (1981). *Climate change and society: consequences of increasing atmospheric carbon dioxide* (No. DOE/EV/10281-T3). Aspen Inst. for Humanistic Studies, CO (USA).

Keller, K.L. (1993), "Conceptualizing, measuring and managing customer-based brand equity", *Journal of Marketing*, Vol. 57, January, pp. 1-22.

Kimura, A., Wada, Y., Kamada, A., Masuda, T., Okamoto, M., Goto, S. I., Tsuzuki, D., Cai, D., Oka, T., and Dan, I. (2010). Interactive effects of carbon footprint information and its accessibility on value and subjective qualities of food products. *Appetite*, 55(2), 271-278.

King, S., Dyball, M., Webster, T., Sharpe, A., Worley, A., DeWitt, J., & Jopson, A. (2009). Exploring public attitudes to climate change and travel choices: Deliberative research. *Final Report for Department for Transport*.

Kolk, A., Levy, D., & Pinkse, J. (2008). Corporate responses in an emerging climate regime: the institutionalization and commensuration of carbon disclosure. *European Accounting Review*, 17(4), 719-745.

Kolk, A., & Pinkse, J. (2004). Market strategies for climate change. *European Management Journal*, 22(3), 304-314.

Kressmann, F., Sirgy, M. J., Herrmann, A., Huber, F., Huber, S., & Lee, D. J. (2006). Direct and indirect effects of self-image congruence on brand loyalty. *Journal of Business Research*, 59(9), 955-964.

Lewis, M. A., & Neighbors, C. (2006). Social norms approaches using descriptive drinking norms education: A review of the research on personalized normative feedback. *Journal of American College Health*, 54(4), 213-218.

Lyndhurst, B. (2007). Public understanding of sustainable energy consumption in the home. *Final Report to the Department for Environment Food and Rural Affairs. London, DEFRA.*

Martin, B. A. S., Lang, B., & Wong, S. (2003). "Conclusion Explicitness in Advertising: The Role of Need for Cognition and Argument Quality on Persuasion," *Journal of Advertising*, 32 (4), 57-65.

Minton, A. P., & Rose, R. L. (1997). The effects of environmental concern on environmentally friendly consumer behavior: An exploratory study. *Journal of Business Research*, 40(1), 37-48.

Monroe, K. B. (1976). The influence of price differences and brand familiarity on brand preferences. *Journal of Consumer Research*, 42-49.

Næs, T., & Mevik, B. H. (2001). Understanding the collinearity problem in regression and discriminant analysis. *Journal of Chemometrics*, 15(4), 413-426.

Olofdotter, M., & Juul, J. (2008), Climate Change and the Food Industry – Climate Labelling for Food Products: Potential and Limitations, *Øresund Food Network*, Copenhagen.

Patchen, M. (1970). Participation, achievement, and involvement on the job. (Englewood Cliffs, N. J.: Prentice-Hall, 1970).

Persson, N. (2010). An exploratory investigation of the elements of B2B brand image and its relationship to price premium. *Industrial Marketing Management*, 39(8), 1269-1277.

Pickett-Baker, J., & Ozaki, R. (2008). Pro-environmental products: marketing influence on consumer purchase decision. *Journal of consumer marketing*, 25(5), 281-293.

Rand Corporation (2004), "Consumer power", available at: www.rand.org/scitech/stpi/ourfuture/Consumer/index.html/ (accessed 6 June, 2015).

Rook, D. W., & Fisher, R. J. (1995). Normative influences on impulsive buying behavior. *Journal of consumer research*, 305-313.

Röös, E., & Tjärnemo, H. (2011). Challenges of carbon labelling of food products: a consumer research perspective. *British Food Journal*, 113(8), 982-996.

Sacks, G., Rayner, M., & Swinburn, B. (2009). Impact of front-of-pack 'traffic-light' nutrition labelling on consumer food purchases in the UK. *Health promotion international*, 24(4), 344-352.

Schwartz, S. H. (1973). Normative explanations of helping behavior: A critique, proposal, and empirical test. *Journal of Experimental Social Psychology*, 9(4), 349-364.

Sheeran, P. (2002). Intention—behavior relations: A conceptual and empirical review. *European review of social psychology*, 12(1), 1-36.

Stern, P. C. (2000). New environmental theories: toward a coherent theory of environmentally significant behavior. *Journal of social issues*, 56(3), 407-424.

Stern, P. C., Dietz, T., & Guagnano, G. A. (1995). The new ecological paradigm in social-psychological context. *Environment and behavior*, 27(6), 723-743.

Strasser, A. A., Tang, K. Z., Romer, D., Jepson, C., & Cappella, J. N. (2012). Graphic warning labels in cigarette advertisements: recall and viewing patterns. *American journal of preventive medicine*, 43(1), 41-47.

Thøgersen, J. (2000). Psychological determinants of paying attention to eco-labels in purchase decisions: Model development and multinational validation. *Journal of Consumer Policy*, 23(3), 285-313.

Tyler, T. R., Orwin, R., & Schurer, L. (1982). Defensive denial and high cost prosocial behavior. *Basic and Applied Social Psychology*, 3(4), 267-281.

Upham, P., & Bleda, M. (2009). Carbon labelling: Public perceptions of the debate. *Tyndall Centre Manchester, The Sustainable Consumption Institute, The University of Manchester: Manchester.*

Upham, P., Dendler, L., & Bleda, M. (2011). Carbon labelling of grocery products: public perceptions and potential emissions reductions. *Journal of Cleaner Production*, 19(4), 348-355.

Valor, C. (2008). Can consumers buy responsibly? Analysis and solutions for market failures. *Journal of Consumer Policy*, 31, 315–326.

van Herpen, E., Hieke, S., & van Trijp, H. (2014). Inferring product healthfulness from nutrition labelling. The influence of reference points. *Appetite*, 72, 138-149.

van Herpen, E., Seiss, E., & van Trijp, H. (2012). The role of familiarity in front-of-pack label evaluation and use: A comparison between the United Kingdom and The Netherlands. *Food Quality and Preference*, 26(1), 22-34.

van Herpen, E., & Trijp, H. (2011). Front-of-pack nutrition labels. Their effect on attention and choices when consumers have varying goals and time constraints. *Appetite*, 57(1), 148-160.

Vanclay, J. K., Shortiss, J., Aulsebrook, S., Gillespie, A. M., Howell, B. C., Johanni, R., & Yates, J. (2011). Customer response to carbon labelling of groceries. *Journal of Consumer Policy*, 34(1), 153-160.

Wee, C. H., Ta, S. J., & Cheok, K. H. (1995). Non-price determinants of intention to purchase counterfeit goods: An exploratory study. *International Marketing Review*, 12(6), 19-46.

Weightman, D., & McDonagh, D. (2004). Supra-functional factors in sustainable products. *Design and manufacture for sustainable development*, 91-101.

Wiidegren, Ö. (1998). The new environmental paradigm and personal norms. *Environment and behavior*, 30(1), 75-100.

Young, W., Hwang, K., McDonald, S., & Oates, C. J. (2010). Sustainable consumption: Green consumer behaviour when purchasing products. *Sustainable Development*, 18, 20–31.

Zintl, R. (1989). Der homo oeconomicus: Ausnahmeerscheinung in jeder Situation oder Jedermann in Ausnahmesituationen. *Analyse & Kritik*, 11(1), 52-69.

Appendices


Appendix I – Types of label

Appendix I.I Checkmark, brand level, positive



Below you see a picture of a refrigerator with some information about different features. Please evaluate this refrigerator based on these features.
Notice that the carbon label is on a brand level, which applies for the brand as a whole and not per s for this specific product.



Net contents (The overall content of the fridge including all sections, measured in litres)	320
Net contents freezer section (The content of the freezer section, measured in litres)	95
Temperature adjustable (Temperature is adjustable both in the regular section as in the freezing section)	Yes
Defrost type (The type of system to defrost with)	Frost gone
Volume (The sound of the refrigerator, measured in decibels)	35 dB
Brand	Brahma
Carbon label brand (The label provides information that indicates whether the brand is carbon friendly or not)	

Appendix I.II Checkmark, brand level, negative



Below you see a picture of a refrigerator with some information about different features. Please evaluate this refrigerator based on these features.
Notice that the carbon label is on a brand level, which applies for the brand as a whole and not per se for this specific product.



Net contents (The overall content of the fridge including all sections, measured in litres)	315
Net contents freezer section (The content of the freezer section, measured in litres)	93
Temperature adjustable (Temperature is adjustable both in the regular section as in the freezing section)	Yes
Defrost type (The type of system to defrost with)	Frost breaker
Volume (The sound of the refrigerator, measured in decibels)	38 dB
Brand	Midea
Carbon label brand (The label provides information that indicates whether the brand is carbon friendly or not)	

Appendix I.III Checkmark, product level, positive



Below you see a picture of a refrigerator with some information about different features. Please evaluate this refrigerator based on these features.

Notice that the carbon label is on a product level, which applies for this specific product only.



Net contents (The overall content of the fridge including all sections, measured in litres)	340
Net contents freezer section (The content of the freezer section, measured in litres)	108
Temperature adjustable (Temperature is adjustable both in the regular section as in the freezing section)	Yes
Defrost type (The type of system to defrost with)	Frost guard
Volume (The sound of the refrigerator, measured in decibels)	39 dB
Brand	YPF
Carbon label (The label provides information that indicates whether a product is produced carbon friendly or not)	


Appendix I.IV Checkmark, product level, negative



Below you see a picture of a refrigerator with some information about different features. Please evaluate this refrigerator **based** on these features.

Notice that the carbon label is on a product level, which applies for this specific product only.




Net contents (The overall content of the fridge including all sections, measured in litres)	337
Net contents freezer section (The content of the freezer section, measured in litres)	110
Temperature adjustable (Temperature is adjustable both in the regular section as in the freezing section)	Yes
Defrost type (The type of system to defrost with)	Frost protector
Volume (The sound of the refrigerator, measured in decibels)	40 dB
Brand	Lan
Carbon label (The label provides information that indicates whether a product is produced carbon friendly or not)	

Appendix I.V Traffic light, brand level, positive



Below you see a picture of a refrigerator with some information about different features. Please evaluate this refrigerator based on these features.
Notice that the carbon label is on a brand level, which applies for the brand as a whole and not per se for this specific product.




Net contents (The overall content of the fridge including all sections, measured in litres)	320
Net contents freezer section (The content of the freezer section, measured in litres)	95
Temperature adjustable (Temperature is adjustable both in the regular section as in the freezing section)	Yes
Defrost type (The type of system to defrost with)	Frost gone
Volume (The sound of the refrigerator, measured in decibels)	35 dB
Brand	Brahma
Carbon label brand (The label provides information that indicates whether the brand is carbon friendly or not)	

Appendix I.VI Traffic light, brand level, negative



Below you see a picture of a refrigerator with some information about different features. Please evaluate this refrigerator based on these features.
Notice that the carbon label is on a brand level, which applies for the brand as a whole and not per se for this specific product.



Net contents (The overall content of the fridge including all sections, measured in litres)	315
Net contents freezer section (The content of the freezer section, measured in litres)	93
Temperature adjustable (Temperature is adjustable both in the regular section as in the freezing section)	Yes
Defrost type (The type of system to defrost with)	Frost breaker
Volume (The sound of the refrigerator, measured in decibels)	38 dB
Brand	Midea
Carbon label brand (The label provides information that indicates whether the brand is carbon friendly or not)	

Appendix I.VII Traffic light, product level, positive



Below you see a picture of a refrigerator with some information about different features. Please evaluate this refrigerator based on these features.

Notice that the carbon label is on a product level, which applies for this specific product only.



Net contents (The overall content of the fridge including all sections, measured in litres)	340
Net contents freezer section (The content of the freezer section, measured in litres)	108
Temperature adjustable (Temperature is adjustable both in the regular section as in the freezing section)	Yes
Defrost type (The type of system to defrost with)	Frost guard
Volume (The sound of the refrigerator, measured in decibels)	39 dB
Brand	YPF
Carbon label (The label provides information that indicates whether a product is produced carbon friendly or not)	


Appendix I.VIII Traffic light, product level, negative



Below you see a picture of a refrigerator with some information about different features. Please evaluate this refrigerator based on these features.

Notice that the carbon label is on a product level, which applies for this specific product only.



Net contents (The overall content of the fridge including all sections, measured in litres)	337
Net contents freezer section (The content of the freezer section, measured in litres)	110
Temperature adjustable (Temperature is adjustable both in the regular section as in the freezing section)	Yes
Defrost type (The type of system to defrost with)	Frost protector
Volume (The sound of the refrigerator, measured in decibels)	40 dB
Brand	Lan
Carbon label (The label provides information that indicates whether a product is produced carbon friendly or not)	

Appendix I.IX Black score, brand level, positive



Below you see a picture of a refrigerator with some information about different features. Please evaluate this refrigerator based on these features.
Notice that the carbon label is on a brand level, which applies for the brand as a whole and not per se for this specific product.



Net contents (The overall content of the fridge including all sections, measured in litres)	320
Net contents freezer section (The content of the freezer section, measured in litres)	95
Temperature adjustable (Temperature is adjustable both in the regular section as in the freezing section)	Yes
Defrost type (The type of system to defrost with)	Frost gone
Volume (The sound of the refrigerator, measured in decibels)	35 dB
Brand	Brahma
Carbon label brand (The label provides information that indicates whether the brand is carbon friendly or not)	87

Appendix I.X Black score, brand level, negative



Below you see a picture of a refrigerator with some information about different features. Please evaluate this refrigerator based on these features.

Notice that the carbon label is on a brand level, which applies for the brand as a whole and not per se for this specific product.



Net contents (The overall content of the fridge including all sections, measured in litres)	315
Net contents freezer section (The content of the freezer section, measured in litres)	93
Temperature adjustable (Temperature is adjustable both in the regular section as in the freezing section)	Yes
Defrost type (The type of system to defrost with)	Frost breaker
Volume (The sound of the refrigerator, measured in decibels)	38 dB
Brand	Midea
Carbon label brand (The label provides information that indicates whether the brand is carbon friendly or not)	29

Appendix I.XI Black score, product level, positive



Below you see a picture of a refrigerator with some information about different features. Please evaluate this refrigerator based on these features.

Notice that the carbon label is on a product level, which applies for this specific product only.



Net contents (The overall content of the fridge including all sections, measured in litres)	340
Net contents freezer section (The content of the freezer section, measured in litres)	108
Temperature adjustable (Temperature is adjustable both in the regular section as in the freezing section)	Yes
Defrost type (The type of system to defrost with)	Frost guard
Volume (The sound of the refrigerator, measured in decibels)	39 dB
Brand	YPF
Carbon label (The label provides information that indicates whether a product is produced carbon friendly or not)	85

Appendix I.XII Black score, product level, negative



Below you see a picture of a refrigerator with some information about different features. Please evaluate this refrigerator based on these features.

Notice that the carbon label is on a product level, which applies for this specific product only.



Net contents (The overall content of the fridge including all sections, measured in litres)	337
Net contents freezer section (The content of the freezer section, measured in litres)	110
Temperature adjustable (Temperature is adjustable both in the regular section as in the freezing section)	Yes
Defrost type (The type of system to defrost with)	Frost protector
Volume (The sound of the refrigerator, measured in decibels)	40 dB
Brand	Lan
Carbon label (The label provides information that indicates whether a product is produced carbon friendly or not)	27

Appendix I.XIII Colored score, brand level, positive



Below you see a picture of a refrigerator with some information about different features. Please evaluate this refrigerator based on these features.

Notice that the carbon label is on a brand level, which applies for the brand as a whole and not per se for this specific product.



Net contents (The overall content of the fridge including all sections, measured in litres)	320
Net contents freezer section (The content of the freezer section, measured in litres)	95
Temperature adjustable (Temperature is adjustable both in the regular section as in the freezing section)	Yes
Defrost type (The type of system to defrost with)	Frost gone
Volume (The sound of the refrigerator, measured in decibels)	35 dB
Brand	Brahma
Carbon label brand (The label provides information that indicates whether the brand is carbon friendly or not)	87

Appendix I.XIV Colored score, brand level, negative



Below you see a picture of a refrigerator with some information about different features. Please evaluate this refrigerator based on these features.
Notice that the carbon label is on a brand level, which applies for the brand as a whole and not per se for this specific product.



Net contents (The overall content of the fridge including all sections, measured in litres)	315
Net contents freezer section (The content of the freezer section, measured in litres)	93
Temperature adjustable (Temperature is adjustable both in the regular section as in the freezing section)	Yes
Defrost type (The type of system to defrost with)	Frost breaker
Volume (The sound of the refrigerator, measured in decibels)	38 dB
Brand	Midea
Carbon label brand (The label provides information that indicates whether the brand is carbon friendly or not)	29

Appendix I.XV Colored score, product level, positive



Below you see a picture of a refrigerator with some information about different features. Please evaluate this refrigerator based on these features.

Notice that the carbon label is on a product level, which applies for this specific product only.



Net contents (The overall content of the fridge including all sections, measured in litres)	340
Net contents freezer section (The content of the freezer section, measured in litres)	108
Temperature adjustable (Temperature is adjustable both in the regular section as in the freezing section)	Yes
Defrost type (The type of system to defrost with)	Frost guard
Volume (The sound of the refrigerator, measured in decibels)	39 dB
Brand	YPF
Carbon label (The label provides information that indicates whether a product is produced carbon friendly or not)	85

Appendix I.XVI Colored score, product level, negative



Below you see a picture of a refrigerator with some information about different features. Please evaluate this refrigerator based on these features.

Notice that the carbon label is on a product level, which applies for this specific product only.



Net contents (The overall content of the fridge including all sections, measured in litres)	337
Net contents freezer section (The content of the freezer section, measured in litres)	110
Temperature adjustable (Temperature is adjustable both in the regular section as in the freezing section)	Yes
Defrost type (The type of system to defrost with)	Frost protector
Volume (The sound of the refrigerator, measured in decibels)	40 dB
Brand	Lan
Carbon label (The label provides information that indicates whether a product is produced carbon friendly or not)	27

Appendix II

Appendix II.I Checkmark label, positive

Carbon label

(The label provides information that indicates whether a product is produced carbon friendly or not)



Appendix II.II checkmark label, negative

Carbon label

(The label provides information that indicates whether a product is produced carbon friendly or not)



Appendix II.III Traffic light label, positive

Carbon label

(The label provides information that indicates whether a product is produced carbon friendly or not)



Appendix II.IV Traffic light label, negative

Carbon label

(The label provides information that indicates whether a product is produced carbon friendly or not)



Appendix II.V Black score label, positive

Carbon label (The label provides information that indicates whether a product is produced carbon friendly or not)	85
---	----

Appendix II.VI Black score label, negative

Carbon label (The label provides information that indicates whether a product is produced carbon friendly or not)	27
---	----

Appendix II.VII Colored score label, positive

Carbon label (The label provides information that indicates whether a product is produced carbon friendly or not)	85
---	----

Appendix II.VIII Colored score label, negative

Carbon label (The label provides information that indicates whether a product is produced carbon friendly or not)	27
---	----

Appendix III – Intro text and manipulation check



Please read this text carefully.

Imagine the following situation:

Recently you have decided that your kitchen needs remodeling. After spending some time making decisions about all of the appliances, the only item that remains to be chosen is a refrigerator. After several hours of searching and comparing you end up with 2 different refrigerators within your budget.

On the next pages you will see two pictures of refrigerators with some of their features. Please evaluate these refrigerators based on their features.

One of the features is a carbon label, which indicates the carbon emission friendliness. Notice that the carbon label is on a brand level, which indicates the score of the brand for all of its products combined. This score does not necessarily apply to this specific refrigerator. The green traffic light indicates that it is a carbon friendly brand, the orange traffic light indicates that the brand is somewhat friendly and unfriendly, while the red traffic light indicates that it is a carbon unfriendly brand.

Please look at the pictures and text carefully and answer the questions.

>>



Below we ask you some questions to see if you understood the situation:

What will you buy in the next question?

- ☐ A refrigerator
- ☐ A television
- ☐ A washing machine

What is the carbon label about?

- ☐ The carbon label provides information about the carbon emission friendliness of this specific product.
- ☐ The carbon label provides information about the carbon emission friendliness of the brand, which applies to all products of this brand combined.

>>

Appendix IV – Negative carbon label brand level

Carbon label brand

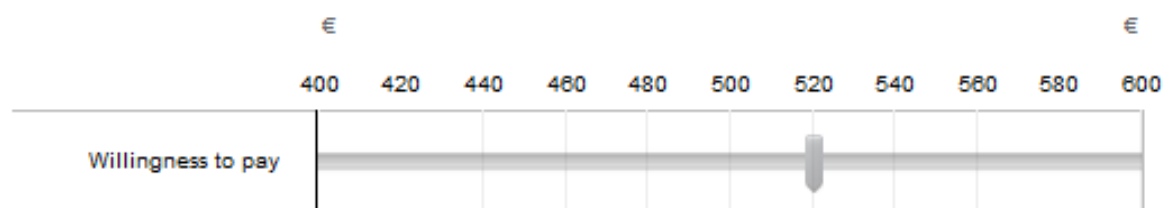
(The label provides information that indicates whether the brand is carbon friendly or not)



Appendix V – WTP, liking and attractiveness

Prices of these fridges vary between €400 and €600.

Please indicate below how much you would be willing to pay for this fridge.



How much do you like this product?



How attractive do you consider this product?



>>


Appendix VI – Negative label product level

Carbon label

(The label provides information that indicates whether a product is produced carbon friendly or not)



Appendix VII – Label assessment



Please state to what extent you agree with the following statements.

	Strongly disagree		Neutral		Strongly agree
I understand the information on the carbon label	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The carbon label gives important information	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The carbon label makes me more aware of the aim of reducing carbon emissions	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The carbon label influences my decision	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The information that the carbon label provides is difficult to interpret	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I take the carbon label seriously	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The carbon label provided me with enough information to estimate the carbon emission friendliness of the brand	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The carbon label provided me with enough information to estimate the carbon emission friendliness of the product	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

How was the information with regard to carbon emissions presented?

☐ A checkmark
☐ A bar
☐ A number
☐ A traffic light
☐ A fuel gage
☐ A color
☐ A colored number

>>

Appendix VIII – Example recall checkmark label



Please look at the picture below, this was the first product that you evaluated. Please indicate the label that was given to this product.



Net contents (The overall content of the fridge including all sections, measured in litres)	320
Net contents freezer section (The content of the freezer section, measured in litres)	95
Temperature adjustable (Temperature is adjustable both in the regular section as in the freezing section)	Yes
Defrost type (The type of system to defrost with)	Frost gone
Volume (The sound of the refrigerator, measured in decibels)	35 dB
Brand	Brahma
Carbon label brand (The label provides information that indicates whether the brand is carbon friendly or not)	



Appendix IX – Example recall traffic light label



Please look at the picture below, this was the first product that you evaluated. Please indicate the label that was given to this product.



Net contents (The overall content of the fridge including all sections, measured in litres)	320
Net contents freezer section (The content of the freezer section, measured in litres)	95
Temperature adjustable (Temperature is adjustable both in the regular section as in the freezing section)	Yes
Defrost type (The type of system to defrost with)	Frost gone
Volume (The sound of the refrigerator, measured in decibels)	35 dB
Brand	Brahma
Carbon label brand (The label provides information that indicates whether the brand is carbon friendly or not)	



Appendix X – Example recall score label



Please look at the picture below, this was the first brand that you evaluated. Please indicate the score of the label that was given to this brand, if you do not know the exact score, please type in your best estimate.



Net contents (The overall content of the fridge including all sections, measured in litres)	320
Net contents freezer section (The content of the freezer section, measured in litres)	95
Temperature adjustable (Temperature is adjustable both in the regular section as in the freezing section)	Yes
Defrost type (The type of system to defrost with)	Frost gone
Volume (The sound of the refrigerator, measured in decibels)	35 dB
Brand	Brahma
Carbon label brand (The label provides information that indicates whether the brand is carbon friendly or not)	


Please indicate the score:

What color was the score?

- ☐ Red
- ☐ Orange
- ☐ Green



Appendix XI – Survey end

 **WAGENINGENUR**
For quality of life

Please fill in the specifics about yourself below

Your gender:

Your age (years):

I am currently completing my:

If you answered 'Other' for the question above, please state your answer:

>>

Debriefing

This is the end of this research. Thank you for your time. Below you can write down your name and e-mail if you want to be considered to make a chance to win an iPad. The lucky winners will be randomly drawn and will receive an e-mail at the end of February.

Purpose

The main purpose of this study is to examine your product evaluations and see how this relates to the tests you filled in before these evaluations.

The main purpose of this study is to examine the relationship between different types of carbon labels and product evaluations and assess how this relates to the tests you filled in before these evaluations. If you are interested in receiving more information about the specifics of the study or the results, please contact the researchers at thijs.boomsma@wur.nl or remco.wezepoel@wur.nl.

Confidentiality

As stated earlier, all your responses will be treated as absolutely confidential. In return, we ask you to honor our confidentiality –please do not tell anyone about the details of this study before March. If the other students know about the purpose of this study before they participate, their data will be biased and thus cannot be included. If this happens it could for example mean that your participation and data become less useful. Please be aware of this and honor our confidentiality.

If you have any questions or comments related to this research (such as how this research can be improved) please write them down here.

To be eligible for the prizes, 2 iPad's, you will have to enter your email address below. If you do not wish to provide your email address you can leave it empty.

Please enter your email address to be able to participate in the lottery for one of the 2 iPad's

Click ">>" to end the research.



Appendix XII – Survey intro and independent variables

Appendix XII.I Introduction



Dear Sir/Madam,

Welcome! We are grateful for your cooperation in this research for our Master thesis about consumer behavior. This research is about evaluating several products on different aspects.

Who is organizing the study?

This research is being conducted by two students from the University of Wageningen, in order to write their Master thesis, under supervision of Dr. M. Handgraaf, Economics of Consumers and Households, Wageningen University.

Instruction and procedure

In this survey you will primarily be asked to what extent you agree or disagree with selected statements. You only need to click on your chosen answer. Instructions for answering the survey questions will be provided to you throughout the survey.

The survey will take less than 20 minutes to complete. For your participation to be successful, it is important you answer all questions.

Voluntary

Your participation is voluntary. If you do not wish to participate in the study, or if you later wish to withdraw your participation, you can do so at any time without penalty.

Confidentiality of the research

Your individual data will only be analysed by the researchers and their supervisor, and will not be provided to third parties. While the aggregate results will be shared, individual results will remain confidential.

Further information

For further information about this study, please contact the researchers: Thijs Boomsma (thijs.boomsma@wur.nl) or Remco Wezepoel (remco.wezepoel@wur.nl).

Thank you in advance for your help. You can now start the questionnaire by clicking the 'continue' button.

Kind regards,

>>

Appendix XII.II Terms of agreement



Before you start with this research it is important to carefully read the following points.

- Your participation in this study is voluntary
- You may withdraw and discontinue participation at any time without penalty
- Your responses will remain anonymous and confidential

I have read and understand the explanation and I:

- ☐ Agree and will continue with the survey
- ☐ Disagree (I will not participate in this survey)



Appendix XII.III Revised NEP scale

Listed below are statements about the relationship between humans and the environment. For each one, please indicate to what extent you agree

	Strongly disagree			Neutral			Strongly agree
We are approaching the limit of the number of people the Earth can support.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Humans have the right to modify the natural environment to suit their needs.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
When humans interfere with nature it often produces disastrous consequences.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Human ingenuity will insure that we do NOT make the Earth unlivable.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Humans are severely abusing the environment.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The earth has plenty of natural resources if we just learn how to develop them.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Plants and animals have as much right as humans to exist.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The balance of nature is strong enough to cope with the impacts to modern industrial nations.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Despite our social abilities humans are still subject to laws of nature.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The so-called 'ecological crisis' facing humankind has been greatly exaggerated.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The earth is like a spaceship with very limited room and resources.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Humans were meant to rule over the rest of nature.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The balance of nature is very delicate and easily upset.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Humans will eventually learn enough about how nature works to be able to control it.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
If things continue on their present course, we will soon experience a major ecological catastrophe.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Appendix XII.IV Perceptions of climate change



The following statements are about climate change.

Please state to what extent you agree with the following statements.

	Strongly disagree			Neutral			Strongly agree
There are simple things that I can do that will have a meaningful effect to alleviate the negative effects of climate change	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I believe that little things I can do will make a difference to alleviate the negative effects of climate change	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I plan to take some actions to stop climate change	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

>>



The following statements are about the global environmental changes.

Please state to what extent you agree with each statement.

	Strongly disagree			Neutral			Strongly agree
I do not feel responsible for this situation	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I think the increase of global environmental changes is not my fault	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
It is not my fault that those environmental changes still increase	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

>>

Listed below are statements about the relationship between humans and the environment. For each one, please indicate to what extent you agree or find them likely

	Very unlikely				Neutral				Very likely
How likely do you think it is that climate change is occurring now?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I have already noticed some signs of climate change	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I am quite sure that human activities are to be blamed for climate change	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The consequences of climate change will be harmful for the environment.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Climate change will bring our world some serious negative consequences	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The consequences of climate change will be more positive than negative.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I intend to take concrete steps to do something to mitigate the negative effects of climate change	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

[>>](#)

Appendix XII.V Norms



Below you will find a sets of questions related to environmental decision making. Again, please indicate to what extent you agree or disagree for each statement.

	Strongly disagree			Neutral			Strongly agree
I feel a moral obligation to protect the environment	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I feel it is important that people in general protect the environment	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Our environmental problems cannot be ignored	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

	Strongly disagree			Neutral			Strongly agree
Most people who are important to me think that purchasing environmentally friendly products or brands is desirable	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Most people who are important to me think that I should buy environmentally friendly products or brands	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I would buy an environment friendly product or brand for the recognition I get from others	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Most people have an interest in environmentally friendly products or brands	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

>>

Appendix XII.VI Need for cognition

Below you see the last set of statements. For each of the statements below, please indicate to what extent the statement is characteristic of you.

	Extremely uncharacteristic				Uncertain				Extremely characteristic
I would prefer complex to simple problems.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I like to have the responsibility of handling a situation that requires a lot of thinking.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Thinking is not my idea of fun.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I would rather do something that requires little thought than something that is sure to challenge my thinking abilities.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I try to anticipate and avoid situations where there is a likely chance I will have to think in-depth about something.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I find satisfaction in deliberating hard and for long hours.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I only think as hard as I have to.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I prefer to think about small, daily projects to long-term ones.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I like tasks that require little thought once I've learned them.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The idea of relying on thought to make my way up to the top appeals to me.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I really enjoy a task that involves coming up with new solutions to the problem.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Learning new ways to think doesn't excite me very much.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I prefer my life to be filled with puzzles that I must solve.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The notion of thinking abstractly is appealing to me.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

I would prefer a task that is intellectual, difficult, and important to one that is somewhat important but does not require much thought.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I feel relief rather than satisfaction after completing a task that required a lot of mental effort.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
It's enough for me that something gets the job done; I don't care how or why it works.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I usually end up deliberating about issues even when they do not affect me personally.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

>>

Appendix XII.VII End first part



You have now completed the first part of the survey. In the next part of the survey you will answer questions about different products.

>>

Appendix XII. Descriptives of label assessment

Descriptive Statistics

	Condition	Mean	Std. Deviation	N
Assessment label-The carbon label gives important information	Checkmark	5,00	1,047	32
	Traffic light	5,24	1,182	34
	Black score	4,70	1,447	33
	Colored score	5,22	,975	32
	Total	5,04	1,186	131
Assessment label -The carbon label makes me more aware of the aim of reducing carbon emissions	Checkmark	5,06	1,045	32
	Traffic light	5,32	1,364	34
	Black score	4,88	1,616	33
	Colored score	5,41	,911	32
	Total	5,17	1,272	131
Assessment label -The carbon label influences my decision	Checkmark	5,16	1,347	32
	Traffic light	5,29	1,528	34
	Black score	4,67	1,762	33
	Colored score	5,09	1,118	32
	Total	5,05	1,464	131
Assessment label --The information that the carbon label provides is difficult to interpret	Checkmark	3,75	1,566	32
	Traffic light	4,21	1,805	34
	Black score	4,64	1,454	33
	Colored score	4,06	1,318	32
	Total	4,17	1,565	131
Assessment label -I take the carbon label seriously	Checkmark	4,75	1,344	32
	Traffic light	5,06	1,301	34
	Black score	4,61	1,694	33
	Colored score	4,91	1,027	32
	Total	4,83	1,360	131
Assessment label -I understand the information on the carbon label	Checkmark	5,00	1,414	32
	Traffic light	5,71	,871	34
	Black score	5,27	1,398	33
	Colored score	5,22	1,157	32
	Total	5,31	1,240	131
Assessment label -The carbon label provided me with enough information to estimate the carbon emission friendliness of the brand	Checkmark	3,56	1,605	32
	Traffic light	4,26	1,463	34
	Black score	3,97	1,425	33
	Colored score	4,19	1,281	32
	Total	4,00	1,457	131
Assessment label -The carbon label provided me with enough information to estimate the carbon emission friendliness of the product	Checkmark	4,06	1,777	32
	Traffic light	4,79	1,298	34
	Black score	4,55	1,502	33
	Colored score	5,09	1,228	32
	Total	4,63	1,495	131